GenCore version 5.1.6

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 30, 2004, 18:52:06 ; Search time 94 Seconds
(without alignments)
12132.267 Million cell updates/sec

Title: US-10-051-902A-21
Perfect score: 3614
Sequence: 1 cttacatgtaagctcgtgcc.....aaaaaaaaaaaaaaaaaaa 2017

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1166195 seqs, 282705291 residues

Total number of hits satisfying chosen parameters: 2332390

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10051902/runat_30062004_164725_20527/app_query.fasta_1.2183
-DB=Published_Applications_AA -QFMT=fastan -SUFFIX=rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10051902_@CGN_1_1_13_@runat_30062004_164725_20527
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2570	71.1	510	13	US-10-051-902-22	Sequence 22, Appl
2	2570	71.1	510	13	US-10-051-909-22	Sequence 22, Appl
3	2015	55.8	539	13	US-10-051-902-26	Sequence 26, Appl
4	2015	55.8	539	13	US-10-051-909-26	Sequence 26, Appl
5	1877.5	52.0	529	13	US-10-051-902-28	Sequence 28, Appl
6	1877.5	52.0	529	13	US-10-051-909-28	Sequence 28, Appl
7	1872.5	51.8	513	13	US-10-051-902-20	Sequence 20, Appl
8	1872.5	51.8	513	13	US-10-051-909-20	Sequence 20, Appl
c 9	1823	50.2	364	12	US-10-425-114-48931	Sequence 48931, A
10	1660	45.9	523	13	US-10-051-902-24	Sequence 24, Appl
11	1660	45.9	523	13	US-10-051-909-24	Sequence 24, Appl
12	1589	44.0	548	10	US-09-774-381-40	Sequence 40, Appl
13	1589	44.0	549	13	US-10-051-902-30	Sequence 30, Appl
14	1589	44.0	549	13	US-10-051-909-30	Sequence 30, Appl
15	1582.5	43.8	517	12	US-10-425-114-56035	Sequence 56035, A
16	1546.5	42.8	541	12	US-10-425-114-66733	Sequence 66733, A
17	1502.5	41.6	580	12	US-10-425-114-67056	Sequence 67056, A
18	1437.5	39.8	407	12	US-10-425-114-49353	Sequence 49353, A
19	1427	39.5	546	12	US-10-425-114-63789	Sequence 63789, A
20	1397.5	38.7	424	12	US-10-424-599-199875	Sequence 199875,
21	1366.5	37.8	380	12	US-10-425-114-39509	Sequence 39509, A
22	1357	37.5	488	12	US-10-424-599-169603	Sequence 169603,
23	1357	37.5	502	12	US-10-425-114-51712	Sequence 51712, A
24	1245	34.4	356	12	US-10-425-114-61926	Sequence 61926, A
25	1164	32.2	533	15	US-10-310-154-724	Sequence 724, App
26	1149.5	31.8	327	12	US-10-425-114-68399	Sequence 68399, A
27	1135	31.4	531	12	US-10-424-599-182839	Sequence 182839,
28	1123.5	31.1	408	12	US-10-424-599-199163	Sequence 199163,
29	994.5	27.5	281	12	US-10-425-114-50090	Sequence 50090, A
30	982.5	27.2	417	12	US-10-425-114-49121	Sequence 49121, A
c 31	803	22.1	475	12	US-10-425-114-41750	Sequence 41750, A
32	782.5	21.7	336	12	US-10-425-114-63429	Sequence 63429, A
33	705.5	19.5	249	12	US-10-424-599-190445	Sequence 190445,
34	687.5	19.0	231	12	US-10-425-114-50887	Sequence 50887, A
35	682	18.9	167	13	US-10-051-902-18	Sequence 18, Appl
36	682	18.9	167	13	US-10-051-909-18	Sequence 18, Appl
37	679.5	18.8	487	9	US-09-795-693-27	Sequence 27, Appl
38	679.5	18.8	487	13	US-10-095-139-14	Sequence 14, Appl
39	679.5	18.8	487	14	US-10-156-239-27	Sequence 27, Appl
40	679.5	18.8	487	14	US-10-199-485-27	Sequence 27, Appl
41	679.5	18.8	487	15	US-10-391-399-45	Sequence 45, Appl
42	672	18.6	488	12	US-09-794-822-11	Sequence 11, Appl
43	672	18.6	488	13	US-10-094-059-4	Sequence 4, Appli
44	672	18.6	488	14	US-10-170-528-5	Sequence 5, Appli
45	672	18.6	488	14	US-10-162-012-46	Sequence 46, Appl

ALIGNMENTS

Instant appl.

RESULT 1
US-10-051-902-22
; Sequence 22, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 22
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (102)

US-10-051-902-22

Alignment Scores:

Pred. No.:	8.21e-145	Length:	510
Score:	2570.00	Matches:	510
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	71.11%	Indels:	0
DB:	13	Gaps:	0

US-10-051-902A-21 (1-2017) x US-10-051-902-22 (1-510)

```
Qy         96 ATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTC 155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetAlaSerAlaAlaLeuProGluAlaValAlaProLysLysLysGlyAsnValArgPhe 20

Qy        156 GCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGG 215
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AlaPheAlaCysAlaIleLeuAlaSerMetThrSerIleLeuLeuGlyTyrAspIleGly 40

Qy        216 GTGATGAGCGGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTG 275
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ValMetSerGlyAlaSerLeuTyrIleLysLysAspPheAsnIleSerAspGlyLysVal 60

Qy        276 GAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGGCGG 335
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GluValLeuMetGlyIleLeuAsnLeuTyrSerLeuIleGlySerPheAlaAlaGlyArg 80

Qy        336 ACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCG 395
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ThrSerAspTrpIleGlyArgArgTyrThrIleValPheAlaAlaValIlePhePheAla 100

Qy        396 GGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTG 455
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 Gly***PheLeuMetGlyPheAlaValAsnTyrAlaMetLeuMetPheGlyArgPheVal 120

Qy        456 GCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCG 515
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AlaGlyIleGlyValGlyTyrAlaLeuMetIleAlaProValTyrThrAlaGluValSer 140

Qy        516 CCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATC 575
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 ProAlaSerAlaArgGlyPheLeuThrSerPheProGluValPheIleAsnPheGlyIle 160

Qy        576 CTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGC 635
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 LeuLeuGlyTyrValSerAsnTyrAlaPheSerArgLeuProLeuAsnLeuGlyTrpArg 180

Qy        636 ATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATG 695
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IleMetLeuGlyIleGlyAlaAlaProSerValLeuLeuAlaLeuMetValLeuGlyMet 200

Qy        696 CCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTG 755
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ProGluSerProArgTrpLeuValMetLysGlyArgLeuAlaAspAlaLysValValLeu 220

Qy        756 GAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCC 815
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 GluLysThrSerAspThrAlaGluGluAlaAlaGluArgLeuAlaAspIleLysAlaAla 240

Qy        816 GCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGA 875
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AlaGlyIleProGluGluLeuAspGlyAspValValThrValProLysArgGlySerGly 260

Qy        876 AACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATC 935
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 AsnGluLysArgValTrpLysGluLeuIleLeuSerProThrProAlaMetArgArgIle 280

Qy        936 CTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTC 995
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 LeuLeuSerGlyIleGlyIleHisPhePheGlnHisAlaLeuGlyIleHisSerValVal 300

Qy        996 TTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGC 1055
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 PheTyrSerProLeuValPheLysSerProGlyLeuThrAsnAspLysHisPheLeuGly 320

Qy       1056 ACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATC 1115
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ThrThrTrpProPheGlyValThrLysArgLeuPheIleLeuLeuAlaThrPhePheIle 340

Qy       1116 GACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTC 1175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 AspGlyValGlyArgArgProLeuLeuLeuGlySerThrGlyGlyIleIleLeuSerLeu 360

Qy       1176 ATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTGG 1235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 IleGlyLeuGlyAlaGlyLeuThrValValGlyGlnHisProAspAlaLysIleProTrp 380

Qy       1236 GCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTT 1295
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 AlaIleGlyLeuSerIleAlaSerThrLeuAlaTyrValAlaPhePheSerIleGlyLeu 400

Qy       1296 GGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGC 1355
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 GlyProIleThrTrpValTyrSerSerGluIlePheProLeuGlnValArgAlaLeuGly 420

Qy       1356 TGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTG 1415
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CysSerLeuGlyValAlaAlaAsnArgValThrSerGlyValIleSerMetThrPheLeu 440

Qy       1416 TCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCG 1475
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 SerLeuSerLysAlaIleThrIleGlyGlySerPhePheLeuTyrSerGlyIleAlaAla 460

Qy       1476 CTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAG 1535
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 LeuAlaTrpValPhePheTyrThrTyrLeuProGluThrArgGlyArgThrLeuGluGlu 480

Qy       1536 ATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAGCCAAGGAG 1595
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 MetSerLysLeuPheGlyAspThrAlaAlaAlaSerGluSerAspGluProAlaLysGlu 500

Qy       1596 AAGAAGAAGGTGGAAATGGCCGCCACTAAC 1625
              |||||||||||||||||||||||||||||
Db        501 LysLysLysValGluMetAlaAlaThrAsn 510
```

RESULT 2
US-10-051-909-22
; Sequence 22, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB1163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 22
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (102)
US-10-051-909-22

Alignment Scores:

Pred. No.:	8.21e-145	Length:	510
Score:	2570.00	Matches:	510

```
Percent Similarity:      100.00%        Conservative:  0
Best Local Similarity:   100.00%        Mismatches:    0
Query Match:             71.11%         Indels:        0
DB:                      13             Gaps:          0

US-10-051-902A-21 (1-2017) x US-10-051-909-22 (1-510)

Qy          96 ATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTC 155
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 MetAlaSerAlaAlaLeuProGluAlaValAlaProLysLysLysGlyAsnValArgPhe 20

Qy         156 GCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGG 215
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          21 AlaPheAlaCysAlaIleLeuAlaSerMetThrSerIleLeuLeuGlyTyrAspIleGly 40

Qy         216 GTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTG 275
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          41 ValMetSerGlyAlaSerLeuTyrIleLysLysAspPheAsnIleSerAspGlyLysVal 60

Qy         276 GAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGG 335
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 GluValLeuMetGlyIleLeuAsnLeuTyrSerLeuIleGlySerPheAlaAlaGlyArg 80

Qy         336 ACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCG 395
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          81 ThrSerAspTrpIleGlyArgArgTyrThrIleValPheAlaAlaValIlePhePheAla 100

Qy         396 GGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTG 455
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         101 Gly***PheLeuMetGlyPheAlaValAsnTyrAlaMetLeuMetPheGlyArgPheVal 120

Qy         456 GCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCG 515
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 AlaGlyIleGlyValGlyTyrAlaLeuMetIleAlaProValTyrThrAlaGluValSer 140

Qy         516 CCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATC 575
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         141 ProAlaSerAlaArgGlyPheLeuThrSerPheProGluValPheIleAsnPheGlyIle 160

Qy         576 CTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGC 635
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         161 LeuLeuGlyTyrValSerAsnTyrAlaPheSerArgLeuProLeuAsnLeuGlyTrpArg 180

Qy         636 ATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATG 695
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 IleMetLeuGlyIleGlyAlaAlaProSerValLeuLeuAlaLeuMetValLeuGlyMet 200

Qy         696 CCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTG 755
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         201 ProGluSerProArgTrpLeuValMetLysGlyArgLeuAlaAspAlaLysValValLeu 220

Qy         756 GAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCC 815
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         221 GluLysThrSerAspThrAlaGluGluAlaAlaGluArgLeuAlaAspIleLysAlaAla 240

Qy         816 GCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGA 875
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 AlaGlyIleProGluGluLeuAspGlyAspValValThrValProLysArgGlySerGly 260

Qy         876 AACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATC 935
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         261 AsnGluLysArgValTrpLysGluLeuIleLeuSerProThrProAlaMetArgArgIle 280

Qy         936 CTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTC 995
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         281 LeuLeuSerGlyIleGlyIleHisPhePheGlnHisAlaLeuGlyIleHisSerValVal 300

Qy         996 TTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGC 1055
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 PheTyrSerProLeuValPheLysSerProGlyLeuThrAsnAspLysHisPheLeuGly 320

Qy        1056 ACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATC 1115
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db         321 ThrThrTrpProPheGlyValThrLysArgLeuPheIleLeuLeuAlaThrPhePheIle 340

Qy        1116 GACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGGATAATCCTCTCCCTC 1175
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         341 AspGlyValGlyArgArgProLeuLeuLeuGlySerThrGlyGlyIleIleLeuSerLeu 360

Qy        1176 ATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTGG 1235
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 IleGlyLeuGlyAlaGlyLeuThrValValGlyGlnHisProAspAlaLysIleProTrp 380

Qy        1236 GCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTT 1295
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         381 AlaIleGlyLeuSerIleAlaSerThrLeuAlaTyrValAlaPhePheSerIleGlyLeu 400

Qy        1296 GGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGC 1355
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         401 GlyProIleThrTrpValTyrSerSerGluIlePheProLeuGlnValArgAlaLeuGly 420

Qy        1356 TGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTG 1415
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 CysSerLeuGlyValAlaAlaAsnArgValThrSerGlyValIleSerMetThrPheLeu 440

Qy        1416 TCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCG 1475
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         441 SerLeuSerLysAlaIleThrIleGlyGlySerPhePheLeuTyrSerGlyIleAlaAla 460

Qy        1476 CTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAG 1535
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         461 LeuAlaTrpValPhePheTyrThrTyrLeuProGluThrArgGlyArgThrLeuGluGlu 480

Qy        1536 ATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAGCCAAGGAG 1595
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 MetSerLysLeuPheGlyAspThrAlaAlaAlaSerGluSerAspGluProAlaLysGlu 500

Qy        1596 AAGAAGAAGGTGGAAATGGCCGCCACTAAC 1625
               ||||||||||||||||||||||||||||||
Db         501 LysLysLysValGluMetAlaAlaThrAsn 510

RESULT 3
US-10-051-902-26
; Sequence 26, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:
;  APPLICANT: Allen, Steve
;  APPLICANT:  Hitz, Bill
;  APPLICANT:  Kinney, Tony
;  APPLICANT:  Tingey, Scott
;  TITLE OF INVENTION: Plant Sugar Transport Proteins
;  FILE REFERENCE: BB-1163
;  CURRENT APPLICATION NUMBER: US/10/051,902
;  CURRENT FILING DATE:  2002-01-17
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
;  PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 26
;   LENGTH: 539
;   TYPE: PRT
;   ORGANISM: Triticum aestivum
US-10-051-902-26

Alignment Scores:
Pred. No.:               8.74e-112      Length:        539
Score:                   2015.00        Matches:       408
Percent Similarity:      84.49%         Conservative:  44
Best Local Similarity:   76.26%         Mismatches:    73
Query Match:             55.76%         Indels:        10
DB:                      13             Gaps:          4

US-10-051-902A-21 (1-2017) x US-10-051-902-26 (1-539)

Qy          48 ACTGTACACGGCCCAGAGCGA----------GCCTCCTCCTCCTCTGCACCACCGGAG 95
               |||      ||||||   |||           |||:::|||      |||
```

```
Db          6 ThrGlnGlyGlyProArgArgHisAsnProGlnAlaAlaSerArgGlyAlaSerSerThr 25

Qy         96 ATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTC 155
              |||       ||||||||||||    ||||||   |||||||||||||||||||||||||
Db         26 MetAspArgAlaAlaLeuProAlaAlaValGluProLysLysLysGlyAsnValArgPhe 45

Qy        156 GCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGG 215
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         46 AlaPheAlaCysAlaIleLeuAlaSerMetThrSerIleLeuLeuGlyTyrAspIleGly 65

Qy        216 GTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTG 275
              ||||||||||||||||||||||||||:::||||||       |||:::|||   ::::::
Db         66 ValMetSerGlyAlaSerLeuTyrIleGlnLysAspLeuLysIleAsnAspThrGlnLeu 85

Qy        276 GAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGGCGG 335
              |||||||||||||||||||||||:::|||||||||||||||||||||||||||||||||||
Db         86 GluValLeuMetGlyIleLeuAsnValTyrSerLeuIleGlySerPheAlaAlaGlyArg 105

Qy        336 ACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCG 395
              |||||||||||||||||||||||:::|||||||||||||||||||||||||||||||||||
Db        106 ThrSerAspTrpIleGlyArgArgPheThrIleValPheAlaAlaValIlePhePheAla 125

Qy        396 GGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTG 455
              |||       :::|||||||||:::|||||||||||||||||||||||||||||||||||
Db        126 GlyAlaLeuIleMetGlyPheSerValAsnTyrAlaMetLeuMetPheGlyArgPheVal 145

Qy        456 GCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCG 515
              ||||||||||||||||||||||||||||||||||||||||||    |||   |||||||||
Db        146 AlaGlyIleGlyValGlyTyrAlaLeuMetIleAlaProValAsnThrGlyGluValSer 165

Qy        516 CCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATC 575
              |||||||||||||||||    ||||||||||||||||||||||||||||||||||||||
Db        166 ProAlaSerAlaArgGlyValLeuThrSerPheProGluValPheIleAsnPheGlyIle 185

Qy        576 CTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGC 635
              ||||||||||||||||||||:::||||||:::||||||   |||   ||||||||||||
Db        186 LeuLeuGlyTyrValSerAsnPheAlaPheAlaArgLeuSerLeuArgLeuGlyTrpArg 205

Qy        636 ATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATG 695
              ||||||||||||||||||    ||||||||||||||||    ||||||||||||||||
Db        206 IleMetLeuGlyIleGlyAlaValProSerValLeuLeuAlaPheMetValLeuGlyMet 225

Qy        696 CCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTG 755
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        226 ProGluSerProArgTrpLeuValMetLysGlyArgLeuAlaAspAlaLysValValLeu 245

Qy        756 GAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCC 815
                 ||||||||||||||    ||||||||||||||||||:::||||||||||||   |||
Db        246 AlaLysThrSerAspThrProGluGluAlaAlaGluArgIleAlaAspIleLysThrAla 265

Qy        816 GCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAG---AGAGGGAGC 872
              ||||||||||       ||||||||||||||||||    |||||||||   :::||||||
Db        266 AlaGlyIleProLeuGlyLeuAspGlyAspValValProValProLysAsnLysGlySer 285

Qy        873 GGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGC 932
                    ||||||||||||    |||:::||||||||||||||||||    |||||||||
Db        286 SerGluGluLysArgValLeuLysAspLeuIleLeuSerProThrIleAlaMetArgHis 305

Qy        933 ATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTC 992
              ||||||::::::||||||||||||||||||||||||||    :::    ||||||   :::|||
Db        306 IleLeuIleAlaGlyIleGlyIleHisPhePheGlnGlnSerSerGlyIleAspAlaVal 325

Qy        993 GTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTG 1052
              |||   ||||||||||||||||||||||||    |||:::|||   |||
Db        326 ValLeuTyrSerProLeuValPheLysSerAlaGlyIleThrGlyAspSerArgLeuArg 345

Qy       1053 GGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTC 1112
              |||||||||          |||   |||       :::||||||||:::||||||||
Db        346 GlyThrThrValAlaValGlyAlaThrAsnThrValPheIleLeuValAlaThrPheLeu 365

Qy       1113 ATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCC 1172
              :::|||   :::   ||||||||||||:::|||   ||||||||||||:::::::::|||
Db        366 LeuAspArgIleArgArgArgProLeuValLeuThrSerThrGlyGlyMetLeuValSer 385

Qy       1173 CTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCT 1232
              |||:::||||||      ||||||||||:::   :::|||||||||   ||||||
Db        386 LeuValGlyLeuAlaThrGlyLeuThrValIleSerArgHisProAspGluLysIleThr 405

Qy       1233 TGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGC 1292
              |||||||||   |||   |||         :::||||||||||||||||||||||||||||
Db        406 TrpAlaIleValLeuCysIlePheCysIleMetAlaTyrValAlaPhePheSerIleGly 425

Qy       1293 CTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTG 1352
              ||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db        426 LeuGlyProIleThrTrpValTyrSerSerGluIlePheProLeuHisValArgAlaLeu 445

Qy       1353 GGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGGTCATCTCCATGACCTTC 1412
              ||||||||||||||||||||   ||||||:::||||||||||||||||||||||||||||
Db        446 GlyCysSerLeuGlyValAlaValAsnArgLeuThrSerGlyValIleSerMetThrPhe 465

Qy       1413 CTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCC 1472
              :::|||||||||||||||:::|||||||||||:::|||||||||::::::|||||||||
Db        466 IleSerLeuSerLysAlaMetThrIleGlyGlyAlaPhePheLeuPheAlaGlyIleAla 485

Qy       1473 GCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAG 1532
              :::   ||||||||||||||:::   ||||||||||||||||||||||||||||||||||
Db        486 SerPheAlaTrpValPhePhePheAlaTyrLeuProGluThrArgGlyArgThrLeuGlu 505

Qy       1533 GAGATGAGCAAGCTGTTCGGCGACACGGCC----------GCCGCCTCGGAATCAGAC 1580
              :::||||||   |||||||||:::||||||              |||:::|||:::|||
Db        506 AspMetSerSerLeuPheGlyAsnThrAlaThrHisLysGlnGlyAlaAlaGluAlaAsp 525

Qy       1581 GAGCCAGCCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCACTAAC 1625
              :::   |||   |||   ||||||||||||||||||||||||
Db        526 AspAspAlaGlyGlu---LysLysValGluMetAlaAlaThrAsn 539

RESULT 4
US-10-051-909-26
; Sequence 26, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
;  APPLICANT: Allen, Steve
;  APPLICANT:  Helentjaris, Tim
;  APPLICANT:  Hitz, Bill
;  APPLICANT:  Kinney, Tony
;  APPLICANT:  Tingey, Scott
;  TITLE OF INVENTION: Plant Sugar Transport Proteins
;  FILE REFERENCE: BB1163 US CIP
;  CURRENT APPLICATION NUMBER: US/10/051,909
;  CURRENT FILING DATE:  2002-01-17
;  PRIOR APPLICATION NUMBER: 60/083,044
;  PRIOR FILING DATE: April 24, 1998
;  NUMBER OF SEQ ID NOS: 38
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 26
;   LENGTH: 539
;   TYPE: PRT
;   ORGANISM: Triticum aestivum
US-10-051-909-26

Alignment Scores:
Pred. No.:                  8.74e-112       Length:          539
Score:                      2015.00         Matches:         408
Percent Similarity:         84.49%          Conservative:    44
Best Local Similarity:      76.26%          Mismatches:      73
Query Match:                55.76%          Indels:          10
DB:                         13              Gaps:            4

US-10-051-902A-21 (1-2017) x US-10-051-909-26 (1-539)

Qy          48 ACTGTACACGGCCCAGAGCGA----------GCCTCCTCCTCCTCTGCACCACCGGAG 95
               |||       ||||||   |||         |||:::|||       |||
Db           6 ThrGlnGlyGlyProArgArgHisAsnProGlnAlaAlaSerArgGlyAlaSerSerThr 25

Qy          96 ATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTC 155
               |||       ||||||||||||    ||||||   |||||||||||||||||||||||||
```

```
Db        26 MetAspArgAlaAlaLeuProAlaAlaValGluProLysLysLysGlyAsnValArgPhe 45

Qy       156 GCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGG 215
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        46 AlaPheAlaCysAlaIleLeuAlaSerMetThrSerIleLeuLeuGlyTyrAspIleGly 65

Qy       216 GTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTG 275
             ||||||||||||||||||||||||||:::||||||      |||:::|||   ::::::
Db        66 ValMetSerGlyAlaSerLeuTyrIleGlnLysAspLeuLysIleAsnAspThrGlnLeu 85

Qy       276 GAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGGCGG 335
             ||||||||||||||||||||||||:::||||||||||||||||||||||||||||||||||
Db        86 GluValLeuMetGlyIleLeuAsnValTyrSerLeuIleGlySerPheAlaAlaGlyArg 105

Qy       336 ACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCG 395
             ||||||||||||||||||||||:::||||||||||||||||||||||||||||||||||||
Db       106 ThrSerAspTrpIleGlyArgArgPheThrIleValPheAlaAlaValIlePhePheAla 125

Qy       396 GGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTG 455
             |||      :::||||||||:::||||||||||||||||||||||||||||||||||||||
Db       126 GlyAlaLeuIleMetGlyPheSerValAsnTyrAlaMetLeuMetPheGlyArgPheVal 145

Qy       456 GCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCG 515
             ||||||||||||||||||||||||||||||||||||    |||   |||||||||
Db       146 AlaGlyIleGlyValGlyTyrAlaLeuMetIleAlaProValAsnThrGlyGluValSer 165

Qy       516 CCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATC 575
             ||||||||||||||||||   ||||||||||||||||||||||||||||||||||||||
Db       166 ProAlaSerAlaArgGlyValLeuThrSerPheProGluValPheIleAsnPheGlyIle 185

Qy       576 CTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGC 635
             ||||||||||||||||||:::|||||||:::||||||   |||   ||||||||||||
Db       186 LeuLeuGlyTyrValSerAsnPheAlaPheAlaArgLeuSerLeuArgLeuGlyTrpArg 205

Qy       636 ATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATG 695
             ||||||||||||||||||   ||||||||||||||||   ||||||||||||||||
Db       206 IleMetLeuGlyIleGlyAlaValProSerValLeuLeuAlaPheMetValLeuGlyMet 225

Qy       696 CCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTG 755
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       226 ProGluSerProArgTrpLeuValMetLysGlyArgLeuAlaAspAlaLysValValLeu 245

Qy       756 GAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCC 815
                ||||||||||||||   ||||||||||||||||||:::||||||||||||   |||
Db       246 AlaLysThrSerAspThrProGluGluAlaAlaGluArgIleAlaAspIleLysThrAla 265

Qy       816 GCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAG---AGAGGGAGC 872
             ||||||||||| |||||||||||||||||| |||||||| :::||||||
Db       266 AlaGlyIleProLeuGlyLeuAspGlyAspValValProValProLysAsnLysGlySer 285

Qy       873 GGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGC 932
                 ||||||||||   |||:::||||||||||||||||||   ||||||||
Db       286 SerGluGluLysArgValLeuLysAspLeuIleLeuSerProThrIleAlaMetArgHis 305

Qy       933 ATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTC 992
             ||||||::::::||||||||||||||||||||||||   :::   ||||||   :::|||
Db       306 IleLeuIleAlaGlyIleGlyIleHisPhePheGlnGlnSerSerGlyIleAspAlaVal 325

Qy       993 GTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTG 1052
             |||   ||||||||||||||||||||||   |||:::|||   |||
Db       326 ValLeuTyrSerProLeuValPheLysSerAlaGlyIleThrGlyAspSerArgLeuArg 345

Qy      1053 GGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTC 1112
             ||||||||          |||   |||      :::|||||||||:::|||||||||
Db       346 GlyThrThrValAlaValGlyAlaThrAsnThrValPheIleLeuValAlaThrPheLeu 365

Qy      1113 ATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCC 1172
             :::|||   :::   ||||||||||||:::|||   ||||||||||:::::::::|||
Db       366 LeuAspArgIleArgArgArgProLeuValLeuThrSerThrGlyGlyMetLeuValSer 385

Qy      1173 CTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCT 1232
             |||:::||||||      ||||||||||||:::   :::|||||||||   ||||||
Db       386 LeuValGlyLeuAlaThrGlyLeuThrValIleSerArgHisProAspGluLysIleThr 405

Qy      1233 TGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGC 1292
             ||||||||   |||   |||         :::|||||||||||||||||||||||||||
Db       406 TrpAlaIleValLeuCysIlePheCysIleMetAlaTyrValAlaPhePheSerIleGly 425

Qy      1293 CTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTG 1352
             ||||||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db       426 LeuGlyProIleThrTrpValTyrSerSerGluIlePheProLeuHisValArgAlaLeu 445

Qy      1353 GGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTC 1412
             ||||||||||||||||||||||||   ||||||:::||||||||||||||||||||||||
Db       446 GlyCysSerLeuGlyValAlaValAsnArgLeuThrSerGlyValIleSerMetThrPhe 465

Qy      1413 CTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCC 1472
             :::||||||||||||||:::||||||||||||:::|||||||||:::::::||||||||
Db       466 IleSerLeuSerLysAlaMetThrIleGlyGlyAlaPhePheLeuPheAlaGlyIleAla 485

Qy      1473 GCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAG 1532
             :::   ||||||||||||||:::   ||||||||||||||||||||||||||||||||||
Db       486 SerPheAlaTrpValPhePhePheAlaTyrLeuProGluThrArgGlyArgThrLeuGlu 505

Qy      1533 GAGATGAGCAAGCTGTTCGGCGACACGGCC----------GCCGCCTCGGAATCAGAC 1580
             :::||||||   |||||||||:::||||||                |||:::|||:::|||
Db       506 AspMetSerSerLeuPheGlyAsnThrAlaThrHisLysGlnGlyAlaAlaGluAlaAsp 525

Qy      1581 GAGCCAGCCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCACTAAC 1625
             :::   |||   |||  ||||||||||||||||||||||||||
Db       526 AspAspAlaGlyGlu---LysLysValGluMetAlaAlaThrAsn 539
```

RESULT 5
US-10-051-902-28
; Sequence 28, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 28
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-051-902-28

Alignment Scores:

Pred. No.:	1.33e-103	Length:	529
Score:	1877.50	Matches:	369
Percent Similarity:	81.38%	Conservative:	55
Best Local Similarity:	70.83%	Mismatches:	92
Query Match:	51.95%	Indels:	5
DB:	13	Gaps:	3

US-10-051-902A-21 (1-2017) x US-10-051-902-28 (1-529)

```
Qy         60 CCAGAGCGA------GCCTCCTCCTCCTCTGCACCACCGGAGATGGCTTCCGCCGCGCTG 113
              |||||||||      |||                    ||||||||||||||||||
Db          5 ProGluArgLysGlyAlaGluAspLysGluGluGlySerArgMetAlaSerAlaAlaLeu 24

Qy        114 CCGGAG------GCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGC 167
              ||||||      ||||||   |||:::   |||||||||   ::::::||||||   |||
Db         25 ProGluProGlyAlaValHisProArgAsnLysGlyAsnPheLysTyrAlaPheThrCys 44

Qy        168 GCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGG 227
              |||:::   |||||||||   :::|||:::|||||||||||||:::|||||||||||||
Db         45 AlaLeuCysAlaSerMetAlaThrIleValLeuGlyTyrAspValGlyValMetSerGly 64
```

```
Qy       228 GCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATG 287
Db        65 AlaSerLeuTyrIleLysArgAspLeuGlnIleThrAspValGlnLeuGluIleMetMet  84

Qy       288 GGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGG 347
Db        85 GlyIleLeuSerValTyrAlaLeuIleGlySerPheLeuGlyAlaArgThrSerAspTrp 104

Qy       348 ATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGSGTTCCTC 407
Db       105 ValGlyArgArgValThrValValPheAlaAlaAlaIlePheAsnAsnGlySerLeuLeu 124

Qy       408 ATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGC 467
Db       125 MetGlyPheAlaValAsnTyrAlaMetLeuMetValGlyArgPheValThrGlyIleGly 144

Qy       468 GTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCG 527
Db       145 ValGlyTyrAlaIleMetValAlaProValTyrThrProGluValSerProAlaSerAla 164

Qy       528 CGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTAC 587
Db       165 ArgGlyPheLeuThrSerPheThrGluValPheIleAsnValGlyIleLeuLeuGlyTyr 184

Qy       588 GTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGC 647
Db       185 ValSerAsnTyrAlaPheAlaArgLeuProLeuHisLeuSerTrpArgValMetLeuGly 204

Qy       648 ATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCG 707
Db       205 IleGlyAlaValProSerAlaLeuLeuAlaLeuMetValPheGlyMetProGluSerPro 224

Qy       708 CGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCC 767
Db       225 ArgTrpLeuValMetLysGlyArgLeuAlaAspAlaArgAlaValLeuAlaLysThrSer 244

Qy       768 GACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCT 827
Db       245 AspThrProGluGluAlaValGluArgLeuAspGlnIleLysAlaAlaAlaGlyIlePro 264

Qy       828 GAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGG 887
Db       265 ArgGluLeuAspGlyAspValValValMetProLysThrLysGlyGlyGlnGluLysGln 284

Qy       888 GTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGG 947
Db       285 ValTrpLysGluLeuIlePheSerProThrProAlaMetArgArgIleLeuLeuAlaAla 304

Qy       948 ATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCT 1007
Db       305 LeuGlyIleHisPhePheGlnGlnAlaThrGlySerAspSerValValLeuTyrSerPro 324

Qy      1008 CTCGTGTTCAAGAGCCCCCGGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCG 1067
Db       325 ArgValPheGlnSerAlaGlyIleThrGlyAspAsnHisLeuLeuGlyAlaThrCysAla 344

Qy      1068 TTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGG 1127
Db       345 MetGlyValMetLysThrLeuPheIleLeuValAlaThrPheGlnLeuAspArgValGly 364

Qy      1128 CGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGC 1187
Db       365 ArgArgProLeuLeuLeuThrSerThrAlaGlyMetLeuAlaCysLeuIleGlyLeuGly 384

Qy      1188 GCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTA 1247
Db       385 ThrGlyLeuThrValValGlyArgHisProAspAlaLysValProTrpAlaIleGlyLeu 404

Qy      1248 AGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACG 1307
Db       405 CysIleValSerIleLeuAlaTyrValSerPhePheSerIleGlyLeuGlyProLeuThr 424

Qy      1308 TGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGC 1367
Db       425 SerValTyrThrSerGluValPheProLeuArgValArgAlaLeuGlyPheAlaLeuGly 444

Qy      1368 GTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAG 1427
Db       445 ThrSerCysAsnArgValThrSerAlaAlaValSerMetSerPheLeuSerLeuSerLys 464

Qy      1428 GCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTG 1487
Db       465 AlaIleThrIleGlyGlySerPhePheLeuTyrAlaGlyIleAlaAlaIleGlyTrpIle 484

Qy      1488 TTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTG 1547
Db       485 PhePhePheThrPheIleProGluThrArgGlyLeuProLeuGluGluIleGlyLysLeu 504

Qy      1548 TTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAGCC---AAGGAGAAGAAGAAG 1604
Db       505 PheGlyMetThrAspThrAlaValGluAlaGlnAspThrAlaThrLysAspLysAlaLys 524

Qy      1605 GTG 1607
Db       525 Val 525
```

RESULT 6
US-10-051-909-28
; Sequence 28, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB1163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 28
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-051-909-28

Alignment Scores:

Pred. No.:	1.33e-103	Length:	529
Score:	1877.50	Matches:	369
Percent Similarity:	81.38%	Conservative:	55
Best Local Similarity:	70.83%	Mismatches:	92
Query Match:	51.95%	Indels:	5
DB:	13	Gaps:	3

US-10-051-902A-21 (1-2017) x US-10-051-909-28 (1-529)

```
Qy         60 CCAGAGCGA------GCCTCCTCCTCCTCTGCACCACCGGAGATGGCTTCCGCCGCGCTG 113
Db          5 ProGluArgLysGlyAlaGluAspLysGluGluGlySerArgMetAlaSerAlaAlaLeu  24

Qy        114 CCGGAG------GCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGC 167
Db         25 ProGluProGlyAlaValHisProArgAsnLysGlyAsnPheLysTyrAlaPheThrCys  44

Qy        168 GCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGG 227
Db         45 AlaLeuCysAlaSerMetAlaThrIleValLeuGlyTyrAspValGlyValMetSerGly  64

Qy        228 GCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATG 287
Db         65 AlaSerLeuTyrIleLysArgAspLeuGlnIleThrAspValGlnLeuGluIleMetMet  84
```

```
Qy        288 GGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGG 347
Db         85 GlyIleLeuSerValTyrAlaLeuIleGlySerPheLeuGlyAlaArgThrSerAspTrp 104

Qy        348 ATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGSGTTCCTC 407
Db        105 ValGlyArgArgValThrValValPheAlaAlaAlaIlePheAsnAsnGlySerLeuLeu 124

Qy        408 ATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGC 467
Db        125 MetGlyPheAlaValAsnTyrAlaMetLeuMetValGlyArgPheValThrGlyIleGly 144

Qy        468 GTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCG 527
Db        145 ValGlyTyrAlaIleMetValAlaProValTyrThrProGluValSerProAlaSerAla 164

Qy        528 CGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTAC 587
Db        165 ArgGlyPheLeuThrSerPheThrGluValPheIleAsnValGlyIleLeuLeuGlyTyr 184

Qy        588 GTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGC 647
Db        185 ValSerAsnTyrAlaPheAlaArgLeuProLeuHisLeuSerTrpArgValMetLeuGly 204

Qy        648 ATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCG 707
Db        205 IleGlyAlaValProSerAlaLeuLeuAlaLeuMetValPheGlyMetProGluSerPro 224

Qy        708 CGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCC 767
Db        225 ArgTrpLeuValMetLysGlyArgLeuAlaAspAlaArgAlaValLeuAlaLysThrSer 244

Qy        768 GACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCT 827
Db        245 AspThrProGluGluAlaValGluArgLeuAspGlnIleLysAlaAlaAlaGlyIlePro 264

Qy        828 GAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGG 887
Db        265 ArgGluLeuAspGlyAspValValValMetProLysThrLysGlyGlyGlnGluLysGln 284

Qy        888 GTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGG 947
Db        285 ValTrpLysGluLeuIlePheSerProThrProAlaMetArgArgIleLeuLeuAlaAla 304

Qy        948 ATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCT 1007
Db        305 LeuGlyIleHisPhePheGlnGlnAlaThrGlySerAspSerValValLeuTyrSerPro 324

Qy       1008 CTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCG 1067
Db        325 ArgValPheGlnSerAlaGlyIleThrGlyAspAsnHisLeuLeuGlyAlaThrCysAla 344

Qy       1068 TTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGG 1127
Db        345 MetGlyValMetLysThrLeuPheIleLeuValAlaThrPheGlnLeuAspArgValGly 364

Qy       1128 CGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGC 1187
Db        365 ArgArgProLeuLeuLeuThrSerThrAlaGlyMetLeuAlaCysLeuIleGlyLeuGly 384

Qy       1188 GCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTA 1247
Db        385 ThrGlyLeuThrValValGlyArgHisProAspAlaLysValProTrpAlaIleGlyLeu 404

Qy       1248 AGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACG 1307
Db        405 CysIleValSerIleLeuAlaTyrValSerPhePheSerIleGlyLeuGlyProLeuThr 424

Qy       1308 TGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGC 1367
Db        425 SerValTyrThrSerGluValPheProLeuArgValArgAlaLeuGlyPheAlaLeuGly 444

Qy       1368 GTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAG 1427
Db        445 ThrSerCysAsnArgValThrSerAlaAlaValSerMetSerPheLeuSerLeuSerLys 464

Qy       1428 GCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTG 1487
Db        465 AlaIleThrIleGlyGlySerPhePheLeuTyrAlaGlyIleAlaAlaIleGlyTrpIle 484

Qy       1488 TTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTG 1547
Db        485 PhePhePheThrPheIleProGluThrArgGlyLeuProLeuGluGluIleGlyLysLeu 504

Qy       1548 TTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAGCC---AAGGAGAAGAAGAAG 1604
Db        505 PheGlyMetThrAspThrAlaValGluAlaGlnAspThrAlaThrLysAspLysAlaLys 524

Qy       1605 GTG 1607
Db        525 Val 525
```

RESULT 7
US-10-051-902-20
; Sequence 20, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 20
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Zea mays
US-10-051-902-20

Alignment Scores:

Pred. No.:	2.62e-103	Length:	513
Score:	1872.50	Matches:	372
Percent Similarity:	84.38%	Conservative:	60
Best Local Similarity:	72.66%	Mismatches:	77
Query Match:	51.81%	Indels:	3
DB:	13	Gaps:	2

US-10-051-902A-21 (1-2017) x US-10-051-902-20 (1-513)

```
Qy          96 ATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTC 155
Db           1 MetAlaSerAspGluLeuAlaLysAlaValGluProArgLysLysGlyAsnValLysTyr 20

Qy         156 GCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGG 215
Db          21 AlaSerIleCysAlaIleLeuAlaSerMetAlaSerValIleLeuGlyTyrAspIleGly 40

Qy         216 GTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTG 275
Db          41 ValMetSerGlyAlaAlaMetTyrIleLysLysAspLeuAsnIleThrAspValGlnLeu 60

Qy         276 GAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGG 335
Db          61 GluIleLeuIleGlyIleLeuSerLeuTyrSerLeuPheGlySerPheAlaGlyAlaArg 80

Qy         336 ACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCG 395
Db          81 ThrSerAspArgIleGlyArgArgLeuThrValValPheAlaAlaValIlePhePheVal 100
```

```
Qy        396 GGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTG 455
Db        101 GlySerLeuLeuMetGlyPheAlaValAsnTyrGlyMetLeuMetAlaGlyArgPheVal 120

Qy        456 GCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCG 515
Db        121 AlaGlyValGlyValGlyTyrGlyGlyMetIleAlaProValTyrThrAlaGluIleSer 140

Qy        516 CCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATC 575
Db        141 ProAlaAlaSerArgGlyPheLeuThrThrPheProGluValPheIleAsnIleGlyIle 160

Qy        576 CTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGC 635
Db        161 LeuLeuGlyTyrLeuSerAsnPheAlaPheAlaArgLeuProLeuHisLeuGlyTrpArg 180

Qy        636 ATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATG 695
Db        181 ValMetLeuAlaIleGlyAlaValProSerGlyLeuLeuAlaLeuLeuValPheCysMet 200

Qy        696 CCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTG 755
Db        201 ProGluSerProArgTrpLeuValLeuLysGlyArgLeuAlaAspAlaArgAlaValLeu 220

Qy        756 GAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCC 815
Db        221 GluLysThrSerAlaThrProGluGluAlaAlaGluArgLeuAlaAspIleLysAlaAla 240

Qy        816 GCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCC---AAGAGAGGGAGC 872
Db        241 AlaGlyIleProLysGlyLeuAspGlyAspValValThrValProGlyLysGluGlnGly 260

Qy        873 GGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGC 932
Db        261 GlyGlyGluLeuGlnValTrpLysLysLeuIleLeuSerProThrProAlaValArgArg 280

Qy        933 ATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTC 992
Db        281 IleLeuLeuSerAlaValGlyLeuHisPhePheGlnGlnAlaSerGlySerAspSerVal 300

Qy        993 GTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTG 1052
Db        301 ValGlnTyrSerAlaArgLeuPheLysSerAlaGlyIleThrAspAspAsnLysLeuLeu 320

Qy       1053 GGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTC 1112
Db        321 GlyValThrCysAlaValGlyValThrLysThrPhePheIleLeuValAlaThrPheLeu 340

Qy       1113 ATCGACGGCGTCGGGCGGCGGCGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCC 1172
Db        341 LeuAspArgAlaGlyArgArgProLeuLeuLeuIleSerThrGlyGlyMetIleValSer 360

Qy       1173 CTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCT 1232
Db        361 LeuIleCysLeuGlySerGlyLeuThrValAlaGlyHisHisProAspThrLysValAla 380

Qy       1233 TGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGC 1292
Db        381 TrpAlaValAlaLeuCysIleAlaSerThrLeuSerTyrIleAlaPhePheSerIleGly 400

Qy       1293 CTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTG 1352
Db        401 LeuGlyProIleThrGlyValTyrThrSerGluIlePheProLeuGlnValArgAlaLeu 420

Qy       1353 GGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTC 1412
Db        421 GlyPheAlaValGlyValAlaSerAsnArgValThrSerAlaValIleSerMetThrPhe 440

Qy       1413 CTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCC 1472
Db        441 LeuSerLeuSerLysAlaIleThrIleGlyGlySerPhePheLeuTyrSerGlyIleAla 460

Qy       1473 GCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAG 1532
Db        461 AlaValAlaTrpValPhePhePheThrCysLeuProGluThrArgGlyArgThrLeuGlu 480

Qy       1533 GAGATGAGCAAGCTGTTCGGC------GACACGGCCGCCGCCTCGGAATCAGACGAGCCA 1586
Db        481 GluMetGlyLysLeuPheGlyMetProAspThrGlyMetAlaGluGluAlaGluAspAla 500

Qy       1587 GCCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCACT 1622
Db        501 AlaAlaLysGluLysValValGluLeuProSerSer 512
```

RESULT 8
US-10-051-909-20
; Sequence 20, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB1163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 20
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Zea mays
US-10-051-909-20

Alignment Scores:

Pred. No.:	2.62e-103	Length:	513
Score:	1872.50	Matches:	372
Percent Similarity:	84.38%	Conservative:	60
Best Local Similarity:	72.66%	Mismatches:	77
Query Match:	51.81%	Indels:	3
DB:	13	Gaps:	2

US-10-051-902A-21 (1-2017) x US-10-051-909-20 (1-513)

```
Qy          96 ATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTC 155
Db           1 MetAlaSerAspGluLeuAlaLysAlaValGluProArgLysLysGlyAsnValLysTyr 20

Qy         156 GCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGG 215
Db          21 AlaSerIleCysAlaIleLeuAlaSerMetAlaSerValIleLeuGlyTyrAspIleGly 40

Qy         216 GTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTG 275
Db          41 ValMetSerGlyAlaAlaMetTyrIleLysLysAspLeuAsnIleThrAspValGlnLeu 60

Qy         276 GAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGG 335
Db          61 GluIleLeuIleGlyIleLeuSerLeuTyrSerLeuPheGlySerPheAlaGlyAlaArg 80

Qy         336 ACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCG 395
Db          81 ThrSerAspArgIleGlyArgArgLeuThrValValPheAlaAlaValIlePhePheVal 100

Qy         396 GGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTG 455
Db         101 GlySerLeuLeuMetGlyPheAlaValAsnTyrGlyMetLeuMetAlaGlyArgPheVal 120

Qy         456 GCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCG 515
Db         121 AlaGlyValGlyValGlyTyrGlyGlyMetIleAlaProValTyrThrAlaGluIleSer 140
```

```
Qy       516 CCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATC 575
Db       141 ProAlaAlaSerArgGlyPheLeuThrThrPheProGluValPheIleAsnIleGlyIle 160

Qy       576 CTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGC 635
Db       161 LeuLeuGlyTyrLeuSerAsnPheAlaPheAlaArgLeuProLeuHisLeuGlyTrpArg 180

Qy       636 ATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATG 695
Db       181 ValMetLeuAlaIleGlyAlaValProSerGlyLeuLeuAlaLeuLeuValPheCysMet 200

Qy       696 CCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTG 755
Db       201 ProGluSerProArgTrpLeuValLeuLysGlyArgLeuAlaAspAlaArgAlaValLeu 220

Qy       756 GAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCC 815
Db       221 GluLysThrSerAlaThrProGluGluAlaAlaGluArgLeuAlaAspIleLysAlaAla 240

Qy       816 GCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCC---AAGAGAGGGAGC 872
Db       241 AlaGlyIleProLysGlyLeuAspGlyAspValValThrValProGlyLysGluGlnGly 260

Qy       873 GGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGC 932
Db       261 GlyGlyGluLeuGlnValTrpLysLysLeuIleLeuSerProThrProAlaValArgArg 280

Qy       933 ATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTC 992
Db       281 IleLeuLeuSerAlaValGlyLeuHisPhePheGlnGlnAlaSerGlySerAspSerVal 300

Qy       993 GTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTG 1052
Db       301 ValGlnTyrSerAlaArgLeuPheLysSerAlaGlyIleThrAspAspAsnLysLeuLeu 320

Qy      1053 GGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTC 1112
Db       321 GlyValThrCysAlaValGlyValThrLysThrPhePheIleLeuValAlaThrPheLeu 340

Qy      1113 ATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCC 1172
Db       341 LeuAspArgAlaGlyArgArgProLeuLeuLeuIleSerThrGlyGlyMetIleValSer 360

Qy      1173 CTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCT 1232
Db       361 LeuIleCysLeuGlySerGlyLeuThrValAlaGlyHisHisProAspThrLysValAla 380

Qy      1233 TGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGC 1292
Db       381 TrpAlaValAlaLeuCysIleAlaSerThrLeuSerTyrIleAlaPhePheSerIleGly 400

Qy      1293 CTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTG 1352
Db       401 LeuGlyProIleThrGlyValTyrThrSerGluIlePheProLeuGlnValArgAlaLeu 420

Qy      1353 GGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTC 1412
Db       421 GlyPheAlaValGlyValAlaSerAsnArgValThrSerAlaValIleSerMetThrPhe 440

Qy      1413 CTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCC 1472
Db       441 LeuSerLeuSerLysAlaIleThrIleGlyGlySerPhePheLeuTyrSerGlyIleAla 460

Qy      1473 GCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAG 1532
Db       461 AlaValAlaTrpValPhePhePheThrCysLeuProGluThrArgGlyArgThrLeuGlu 480

Qy      1533 GAGATGAGCAAGCTGTTCGGC------GACACGGCCGCCGCCTCGGAATCAGACGAGCCA 1586
Db       481 GluMetGlyLysLeuPheGlyMetProAspThrGlyMetAlaGluGluAlaGluAspAla 500

Qy      1587 GCCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCACT 1622
```

```
Db       501 AlaAlaLysGluLysValValGluLeuProSerSer 512

RESULT 9
US-10-425-114-48931
; Sequence 48931, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 48931
;   LENGTH: 364
;   TYPE: PRT
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: UC-OSROCYP002H08_FLI.pep
US-10-425-114-48931

Alignment Scores:
Pred. No.:                2.22e-100      Length:          364
Score:                    1823.00        Matches:         341
Percent Similarity:       96.15%         Conservative:    9
Best Local Similarity:    93.68%         Mismatches:      14
Query Match:              50.21%         Indels:          0
DB:                       12             Gaps:            0

US-10-051-902A-21 (1-2017) x US-10-425-114-48931 (1-364)

Qy      1127 CCCGACGCCGTCGATGAAGAAAGTCGCCAACAAGATGAAAAGCCTCTTGGTGACACCGAA 1068
Db          1 ProAspAlaValGluGluGluGlyArgHisGluAspGluGluArgLeuGlyAspThrAsp 20

Qy      1067 CGGCCAAGTGGTGCCCAAGAAGTGTTTGTCGTTCGTTAATCCGGGGCTCTTGAACACGAG 1008
Db         21 GlyAlaGlyGlyAlaGluGluValPheValValGlyAspAlaGlyAlaLeuGluHisAla 40

Qy      1007 AGGGCTGTAGAAGACGACGGAGTGAATGCCCAACGCATGCTGGAAGAAGTGGATGCCGAT 948
Db         41 ArgAlaValGluHisAspGlyValAspAlaArgArgLeuLeuGluGluValAspAlaAsp 60

Qy       947 CCCGGACAGCAGGATGCGCCGCATGGCCGGGGGTCGGGGACAGGATGAGCTCCTTCCACAC 888
Db         61 AlaArgGlnGlnAspAlaProHisGlyArgGlyArgGlyGlnAspGluLeuLeuProHis 80

Qy       887 CCGCTTCTCGTTTCCGCTCCCTCTCTTGGGGACGGTCACCACGTCGCCGTCGAGCTCCTC 828
Db         81 ProLeuLeuValSerAlaProSerLeuGlyAspGlyHisHisValAlaValGluLeuLeu 100

Qy       827 AGGGATGCCGGCGGCGGCCTTGATGTCGGCCAGGCGCTCCGCGGCCTCCTCCGCCGTGTC 768
Db        101 GlyAspAlaGlyGlyGlyLeuAspValGlyGlnAlaLeuArgGlyLeuLeuArgArgVal 120

Qy       767 GGAGGTCTTCTCCAGCACCACCTTGGCGTCCGCGAGGCGTCCCTTCATGACCAGCCACCG 708
Db        121 GlyGlyLeuLeuGlnHisHisLeuGlyValArgGluAlaSerLeuHisAspGlnProPro 140

Qy       707 CGGCGACTCCGGCATGCCGAGCACCATGAGCGCGAGCAGCACGGACGGCGCCGCGCCGAT 648
Db        141 ArgArgLeuArgHisAlaGluHisHisGluArgGluGlnHisGlyArgArgArgAlaAsp 160

Qy       647 GCCGAGCATGATGCGCCACCCGAGGTTCAGCGGCAAGCGGGAGAAAGCATAGTTCGAGAC 588
Db        161 AlaGluHisAspAlaProProGluValGlnArgGlnAlaGlyGluSerIleValArgAsp 180
```

```
Qy        587 GTACCCGAGCAGGATGCCGAAGTTGATGAACACCTCCGGGAACGACGTCAGGAAGCCACG 528
Db        181 ValProGluGlnAspAlaGluValAspGluHisLeuArgGluArgArgGlnGluAlaThr 200

Qy        527 CGCCGACGCCGGCGACACCTCGGCGGTGTACACCGGCGCGATCATGAGCGCGTAGCCCAC 468
Db        201 ArgArgArgArgArgHisLeuGlyGlyValHisArgArgAspHisGluArgValAlaHis 220

Qy        467 GCCGATGCCGGCCACGAAGCGGCCGAACATGAGCATGGCGTAGTTGACGGCGAACCCCAT 408
Db        221 AlaAspAlaGlyHisGluAlaAlaGluHisGluHisGlyValValAspGlyGluProHis 240

Qy        407 GAGGAACSCCCCCGCGAAGAATATGACGGCGGCGAACACGATGGTGTACCGCCGGCCGAT 348
Db        241 GluGluArgProArgGluGluTyrAspGlyGlyGluHisAspGlyValProProAlaAsp 260

Qy        347 CCAGTCCGACGTCCGCCCCGCCGCGAAGGAGCCGATGAGCGAGTAGAGGTTCAGTATGCC 288
Db        261 ProValArgArgProProArgArgGluGlyAlaAspGluArgValGluValGlnTyrAla 280

Qy        287 CATGAGAACCTCCACCTTCCCGTCACTGATGTTGAAGTCCTTCTTGATGTACAGCGACGC 228
Db        281 HisGluAsnLeuHisLeuProValThrAspValGluValLeuLeuAspValGlnArgArg 300

Qy        227 CCCGCTCATCACCCCGATATCGTAGCCGAGGAGGATGGAGGTCATGGAGGCGAGGATGGC 168
Db        301 ProAlaHisHisProAspIleValAlaGluGluAspGlyGlyHisGlyGlyGluAspGly 320

Qy        167 GCAGGCGAAGGCGAACCGGACGTTGCCCTTCTTCTTCGGCGCGACGGCCTCCGGCAGCGC 108
Db        321 AlaGlyGluGlyGluProAspValAlaLeuLeuLeuArgArgAspGlyLeuArgGlnArg 340

Qy        107 GGCGGAAGCCATCTCCGGTGGTGCAGAGGAGGAGGAGGCTCGCTCTGGGCCGTGTACAGT 48
Db        341 GlyGlySerHisLeuArgTrpCysArgGlyGlyGlyGlySerLeuTrpAlaValTyrSer 360

Qy         47 AGTGGCGGTCGA 36
Db        361 SerGlyGlyArg 364
```

RESULT 10
US-10-051-902-24
; Sequence 24, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 24
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Glycine max
US-10-051-902-24

Alignment Scores:

Pred. No.:	1.16e-90	Length:	523
Score:	1660.00	Matches:	331
Percent Similarity:	75.76%	Conservative:	69
Best Local Similarity:	62.69%	Mismatches:	110
Query Match:	45.93%	Indels:	18
DB:	13	Gaps:	4

US-10-051-902A-21 (1-2017) x US-10-051-902-24 (1-523)

```
Qy         99 GCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCC 158
Db         12 AlaHisLysThrLeuGlnAspPheAspProProLysLysArgLysArgAsnLysTyrAla 31

Qy        159 TTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTG 218
Db         32 PheAlaCysAlaMetLeuAlaSerMetThrSerIleLeuLeuGlyTyrAspIleGlyVal 51

Qy        219 ATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAG 278
Db         52 MetSerGlyAlaAlaIleTyrIleLysArgAspLeuLysValSerAspGluGlnIleGlu 71

Qy        279 GTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACG 338
Db         72 IleLeuLeuGlyIleIleAsnLeuTyrSerLeuIleGlySerCysLeuAlaGlyArgThr 91

Qy        339 TCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGG 398
Db         92 SerAspTrpIleGlyProArgTyrThrIleValPheAlaGlyThrIlePhePheValGly 111

Qy        399 GSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCC 458
Db        112 AlaLeuLeuMetGlyPheSerProAsnTyrSerPheLeuMetPheGlyArgPheValAla 131

Qy        459 GGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCG 518
Db        132 GlyIleGlyIleGlyTyrAlaLeuMetIleAlaProValTyrThrAlaGluValSerPro 151

Qy        519 GCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTG 578
Db        152 AlaSerSerArgGlyPheLeuThrSerProGluValPheIleAsnGlyGlyIleLeu 171

Qy        579 CTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATC 638
Db        172 IleGlyTyrIleSerAsnTyrAlaPheSerLysLeuThrLeuLysValGlyTrpArgMet 191

Qy        639 ATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCG 698
Db        192 MetLeuGlyValGlyAlaIleProSerValLeuLeuThrValGlyValLeuAlaMetPro 211

Qy        699 GAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAG 758
Db        212 GluSerProArgTrpLeuValMetArgGlyArgLeuGlyGluAlaArgLysValLeuAsn 231

Qy        759 AAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCC 818
Db        232 LysThrSerAspSerLysGluGluAlaGlnLeuArgLeuAlaGluIleLysGlnAlaAla 251

Qy        819 GGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAAC 878
Db        252 GlyIleProGluSerCysAsnAspAspValValGlnValAsnLysGlnSerAsnGly--- 270

Qy        879 GAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTG 938
Db        271 ---GluGlyValTrpLysGluLeuPheLeuTyrProThrProAlaIleArgHisIleVal 289

Qy        939 CTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTC 998
Db        290 IleAlaAlaLeuGlyIleHisPhePheGlnGlnAlaSerGlyValAspAlaValValLeu 309

Qy        999 TACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGCACC 1058
Db        310 TyrSerProArgIlePheGluLysAlaGlyIleThrAsnAspThrHisLysLeuLeuAla 329

Qy       1059 ACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGAC 1118
Db        330 ThrValAlaValGlyPheValLysThrValPheIleLeuAlaAlaThrPheThrLeuAsp 349

Qy       1119 GGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATC 1178
Db        350 ArgValGlyArgArgProLeuLeuLeuSerSerValGlyGlyMetValLeuSerLeuLeu 369
```

```
Qy       1179 GGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCC 1238
Db        370 ThrLeuAlaIleSerLeuThrValIle---AspHisSerGluArgLysLeuMetTrpAla 388

Qy       1239 ATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGC 1298
Db        389 ValGlySerSerIleAlaMetValLeuAlaTyrValAlaThrPheSerIleGlyAlaGly 408

Qy       1299 CCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGC 1358
Db        409 ProIleThrTrpValTyrSerSerGluIlePheProLeuArgLeuArgAlaGlnGlyAla 428

Qy       1359 TCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCG 1418
Db        429 AlaAlaGlyValAlaValAsnArgThrThrSerAlaValValSerMetThrPheLeuSer 448

Qy       1419 CTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTC 1478
Db        449 LeuThrArgAlaIleThrIleGlyGlyAlaPhePheLeuTyrCysGlyIleAlaThrVal 468

Qy       1479 GCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATG 1538
Db        469 GlyTrpIlePhePheTyrThrValLeuProGluThrArgGlyLysThrLeuGluAspMet 488

Qy       1539 AGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAGCCAAGGAGAAG 1598
Db        489 GluGlySerPheGlyThrPheArgSerLysSerAsnAla----------------Ser 502

Qy       1599 AAGAAGGTGGAAATGGCCGCCACTAACTGATCAAACTAACCGCAAAATCACCAAATCCTA 1658
Db        503 LysAlaValGlu--------------------------AsnGluAsnGlyGlnValAla 513

Qy       1659 AGGGTTTTCTTGCAAAAACGTGTG 1682
Db        514 GlnValGlnLeuGlyThrAsnVal 521

RESULT 11
US-10-051-909-24
; Sequence 24, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
;  APPLICANT: Allen, Steve
;  APPLICANT:  Helentjaris, Tim
;  APPLICANT:  Hitz, Bill
;  APPLICANT:  Kinney, Tony
;  APPLICANT:  Tingey, Scott
;  TITLE OF INVENTION: Plant Sugar Transport Proteins
;  FILE REFERENCE: BB1163 US CIP
;  CURRENT APPLICATION NUMBER: US/10/051,909
;  CURRENT FILING DATE:  2002-01-17
;  PRIOR APPLICATION NUMBER: 60/083,044
;  PRIOR FILING DATE: April 24, 1998
;  NUMBER OF SEQ ID NOS: 38
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 24
;   LENGTH: 523
;   TYPE: PRT
;   ORGANISM: Glycine max
US-10-051-909-24

Alignment Scores:
Pred. No.:                1.16e-90      Length:          523
Score:                    1660.00       Matches:         331
Percent Similarity:       75.76%        Conservative:    69
Best Local Similarity:    62.69%        Mismatches:      110
Query Match:              45.93%        Indels:          18
DB:                       13            Gaps:            4

US-10-051-902A-21 (1-2017) x US-10-051-909-24 (1-523)

Qy         99 GCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCC 158
Db         12 AlaHisLysThrLeuGlnAspPheAspProProLysLysArgLysArgAsnLysTyrAla 31

Qy        159 TTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTG 218
Db         32 PheAlaCysAlaMetLeuAlaSerMetThrSerIleLeuLeuGlyTyrAspIleGlyVal 51

Qy        219 ATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAG 278
Db         52 MetSerGlyAlaAlaIleTyrIleLysArgAspLeuLysValSerAspGluGlnIleGlu 71

Qy        279 GTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACG 338
Db         72 IleLeuLeuGlyIleIleAsnLeuTyrSerLeuIleGlySerCysLeuAlaGlyArgThr 91

Qy        339 TCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGG 398
Db         92 SerAspTrpIleGlyProArgTyrThrIleValPheAlaGlyThrIlePhePheValGly 111

Qy        399 GSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCC 458
Db        112 AlaLeuLeuMetGlyPheSerProAsnTyrSerPheLeuMetPheGlyArgPheValAla 131

Qy        459 GGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCG 518
Db        132 GlyIleGlyIleGlyTyrAlaLeuMetIleAlaProValTyrThrAlaGluValSerPro 151

Qy        519 GCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTG 578
Db        152 AlaSerSerArgGlyPheLeuThrSerProProGluValPheIleAsnGlyGlyIleLeu 171

Qy        579 CTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATC 638
Db        172 IleGlyTyrIleSerAsnTyrAlaPheSerLysLeuThrLeuLysValGlyTrpArgMet 191

Qy        639 ATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCG 698
Db        192 MetLeuGlyValGlyAlaIleProSerValLeuLeuThrValGlyValLeuAlaMetPro 211

Qy        699 GAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAG 758
Db        212 GluSerProArgTrpLeuValMetArgGlyArgLeuGlyGluAlaArgLysValLeuAsn 231

Qy        759 AAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCC 818
Db        232 LysThrSerAspSerLysGluGluAlaGlnLeuArgLeuAlaGluIleLysGlnAlaAla 251

Qy        819 GGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAAC 878
Db        252 GlyIleProGluSerCysAsnAspAspValValGlnValAsnLysGlnSerAsnGly--- 270

Qy        879 GAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTG 938
Db        271 ---GluGlyValTrpLysGluLeuPheLeuTyrProThrProAlaIleArgHisIleVal 289

Qy        939 CTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTC 998
Db        290 IleAlaAlaLeuGlyIleHisPhePheGlnGlnAlaSerGlyValAspAlaValValLeu 309

Qy        999 TACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGCACC 1058
Db        310 TyrSerProArgIlePheGluLysAlaGlyIleThrAsnAspThrHisLysLeuLeuAla 329

Qy       1059 ACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGAC 1118
Db        330 ThrValAlaValGlyPheValLysThrValPheIleLeuAlaAlaThrPheThrLeuAsp 349

Qy       1119 GGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATC 1178
Db        350 ArgValGlyArgArgProLeuLeuLeuSerSerValGlyGlyMetValLeuSerLeuLeu 369

Qy       1179 GGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCC 1238
Db        370 ThrLeuAlaIleSerLeuThrValIle---AspHisSerGluArgLysLeuMetTrpAla 388
```

```
Qy       1239 ATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGC 1298
              :::|||   |||||||||      ||||||||||||||   ||||||||||||   |||
Db        389 ValGlySerSerIleAlaMetValLeuAlaTyrValAlaThrPheSerIleGlyAlaGly 408

Qy       1299 CCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGC 1358
              ||||||||||||||||||||||||||||||||||:::::|||||   |||
Db        409 ProIleThrTrpValTyrSerSerGluIlePheProLeuArgLeuArgAlaGlnGlyAla 428

Qy       1359 TCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCG 1418
              :::   |||||||||   ||||||   ||||||   |||:::||||||||||||||||||
Db        429 AlaAlaGlyValAlaValAsnArgThrThrSerAlaValValSerMetThrPheLeuSer 448

Qy       1419 CTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTC 1478
              |||::::::||||||||||||||||||:::||||||||||   ||||||||   :::
Db        449 LeuThrArgAlaIleThrIleGlyGlyAlaPhePheLeuTyrCysGlyIleAlaThrVal 468

Qy       1479 GCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATG 1538
                |||:::||||||||||||   ||||||||||||||||||:::|||||||||:::|||
Db        469 GlyTrpIlePhePheTyrThrValLeuProGluThrArgGlyLysThrLeuGluAspMet 488

Qy       1539 AGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAGCCAAGGAGAAG 1598
                    ||||||         :::   |||   :::
Db        489 GluGlySerPheGlyThrPheArgSerLysSerAsnAla------------------Ser 502

Qy       1599 AAGAAGGTGGAAATGGCCGCCACTAACTGATCAAACTAACCGCAAAATCACCAAATCCTA 1658
              |||   ||||||                              :::|||   |||:::
Db        503 LysAlaValGlu--------------------------AsnGluAsnGlyGlnValAla 513

Qy       1659 AGGGTTTTCTTGCAAAAACGTGTG 1682
              :::|||   |||          |||
Db        514 GlnValGlnLeuGlyThrAsnVal 521

RESULT 12
US-09-774-381-40
; Sequence 40, Application US/09774381
; Publication No. US20030082677A1
; GENERAL INFORMATION:
;  APPLICANT:  Holtzman, Douglas A.
;  APPLICANT:  McCarthy, Sean A.
;  APPLICANT:  Pan, Yang
;  APPLICANT:  Gearing, David P.
;  TITLE OF INVENTION:  NOVEL EDIRF, MTR-1, LSP-1, TAP-1, AND PA-I MOLECULES
;  TITLE OF INVENTION:  AND USES THEREFOR
;  FILE REFERENCE:  MNI-107CP2
;  CURRENT APPLICATION NUMBER: US/09/774,381
;  CURRENT FILING DATE:  2001-01-30
;  PRIOR APPLICATION NUMBER:  08/941,354
;  PRIOR FILING DATE:  1999-09-30
;  PRIOR APPLICATION NUMBER:  09/010,674
;  PRIOR FILING DATE:  1998-01-22
;  PRIOR APPLICATION NUMBER:  60/061,149
;  PRIOR FILING DATE:  1997-10-06
;  PRIOR APPLICATION NUMBER:  09/014,347
;  PRIOR FILING DATE:  1998-01-27
;  PRIOR APPLICATION NUMBER:  60/061,159
;  PRIOR FILING DATE:  1997-10-06
;  PRIOR APPLICATION NUMBER:  09/474,151
;  PRIOR FILING DATE:  2000-12-21
;  PRIOR APPLICATION NUMBER:  09/004,206
;  PRIOR FILING DATE:  1998-01-08
;  PRIOR APPLICATION NUMBER:  60/061,143
;  PRIOR FILING DATE:  1997-10-06
;  PRIOR APPLICATION NUMBER:  09/483,414
;  PRIOR FILING DATE:  2000-01-14
;  PRIOR APPLICATION NUMBER:  09/213,571
;  PRIOR FILING DATE:  1998-12-18
;  PRIOR APPLICATION NUMBER:  08/994,890
;  PRIOR FILING DATE:  1997-12-19
;  NUMBER OF SEQ ID NOS:  59
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 40
;   LENGTH: 548
;   TYPE: PRT
;   ORGANISM: Beta vulgaris
US-09-774-381-40

Alignment Scores:
Pred. No.:                1.95e-86      Length:          548
Score:                    1589.00       Matches:         326
Percent Similarity:       72.23%        Conservative:    59
Best Local Similarity:    61.16%        Mismatches:      126
Query Match:              43.97%        Indels:          22
DB:                       10            Gaps:            5

US-10-051-902A-21 (1-2017) x US-09-774-381-40 (1-548)

Qy         78 TCCTCTGCACCACCGGAGATGGCTTCCGCCGCGCTG----------CCGGAGGCCGTC 125
              |||      ||||||      ||||||      :::           |||
Db          9 SerAspProProProThrThrAlaSerLysValIleAlaAspPheAspProLeuLysLys 28

Qy        126 GCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATG 185
                 ||||||:::            :::||||||||||||||||||   ||||||||||
Db         29 ProProLysArgAsn---------LysPheAlaPheAlaCysAlaThrLeuAlaSerMet 45

Qy        186 ACCTCCATCCTCCTCGGCTACGATATCGGGGGTGATGAGCGGGGCGTCGCTGTACATCAAG 245
              ||||||:::||||||||||||||||||||||||||||||||||||   :::|||:::|||
Db         46 ThrSerValLeuLeuGlyTyrAspIleGlyValMetSerGlyAlaIleIleTyrLeuLys 65

Qy        246 AAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTAC 305
              :::|||::::::||||||||   ::::::   ||||||:::|||||||||||||:::|||
Db         66 GluAspTrpHisIleSerAspThrGlnIleGlyValLeuValGlyIleLeuAsnIleTyr 85

Qy        306 TCGCTCATCGGCTCCTTCGCGGCGGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACC 365
                 |||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 CysLeuPheGlySerPheAlaAlaGlyArgThrSerAspTrpIleGlyArgArgTyrThr 105

Qy        366 ATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAAC 425
              ||||||   |||      |||||||||   |||      ||||||||||||||   |||
Db        106 IleValLeuAlaGlyAlaIlePhePheValGlyAlaLeuLeuMetGlyPheAlaThrAsn 125

Qy        426 TACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATG 485
              ||||||   ||||||   ||||||||||||   ||||||||||||||||||||||||||
Db        126 TyrAlaPheLeuMetValGlyArgPheValThrGlyIleGlyValGlyTyrAlaLeuMet 145

Qy        486 ATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCG 545
              ||||||||||||||||||||||||||||||||||||||:::||||||||||||||||||
Db        146 IleAlaProValTyrThrAlaGluValSerProAlaSerSerArgGlyPheLeuThrSer 165

Qy        546 TTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTC 605
              ||||||||||||||||||||   |||||||||||||||||:::||||||   ||||||
Db        166 PheProGluValPheIleAsnAlaGlyIleLeuLeuGlyTyrIleSerAsnLeuAlaPhe 185

Qy        606 TCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCC 665
              |||   ||||||   :::|||   ||||||   ||||||||||||||||||   ||||||
Db        186 SerSerLeuProThrHisLeuSerTrpArgPheMetLeuGlyIleGlyAlaIleProSer 205

Qy        666 GTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAG 725
              :::   ||||||:::   ||||||   ||||||||||||||||||||||||||||:::
Db        206 IlePheLeuAlaIleGlyValLeuAlaMetProGluSerProArgTrpLeuValMetGln 225

Qy        726 GGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCC 785
              |||||||||   |||||||||   |||:||   :::   ||||||:::   |||||||||
Db        226 GlyArgLeuGlyAspAlaLysLysValLeuAsnArgIleSerAspSerProGluGluAla 245

Qy        786 GCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGAC 845
                 ||||||::::::|||||      ||||||||||||   |||   |||   |||
Db        246 GlnLeuArgLeuSerGluIleLysGlnThrAlaGlyIleProAlaGluCysAspGluAsp 265

Qy        846 GTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATC 905
              :::      |||   |||            :::         ||||||||||||||||
Db        266 IleTyrLysValGluLysThrLysIleLysSerGlyAsnAlaValTrpLysGluLeuPhe 285

Qy        906 CTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTC 965
              :::|||||||||||||:::||||||   :::::::::|||||||||||||||||||||
Db        286 PheAsnProThrProAlaValArgArgAlaValIleAlaGlyIleGlyIleHisPhePhe 305
```

```
Qy        966 CAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCC 1025
Db        306 GlnGlnAlaSerGlyIleAspAlaValValLeuTyrSerProArgIlePheGlnSerAla 325

Qy       1026 GGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGG 1085
Db        326 GlyIleThrAsnAlaArgLysGlnLeuLeuAlaThrValAlaValGlyValValLysThr 345

Qy       1086 CTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTG 1145
Db        346 LeuPheIleLeuValAlaThrPheGlnLeuAspLysTyrGlyArgArgProLeuLeuLeu 365

Qy       1146 GGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTC 1205
Db        366 ThrSerValGlyGlyMetIleIleAlaIleLeuThrLeuAlaMetSerLeuThrValIle 385

Qy       1206 GGCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTC 1265
Db        386 ---AspHisSerHisHisLysIleThrTrpAlaIleAlaLeuCysIleThrMetValCys 404

Qy       1266 GCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAG 1325
Db        405 AlaValValAlaSerPheSerIleGlyLeuGlyProIleThrTrpValTyrSerSerGlu 424

Qy       1326 ATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTC 1385
Db        425 ValPheProLeuArgLeuArgAlaGlnGlyThrSerMetGlyValAlaValAsnArgVal 444

Qy       1386 ACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGC 1445
Db        445 ValSerGlyValIleSerIlePhePheLeuProLeuSerHisLysIleThrThrGlyGly 464

Qy       1446 AGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTC 1505
Db        465 AlaPhePheLeuPheGlyGlyIleAlaIleIleAlaTrpPhePhePheLeuThrPheLeu 484

Qy       1506 CCGGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGAC--------- 1556
Db        485 ProGluThrArgGlyArgThrLeuGluAsnMetHisGluLeuPheGluAspPheArgTrp 504

Qy       1557 ------------ACGGCCGCCGCCTCGGAATCAGACGAGCCAGCC--------------- 1589
Db        505 ArgGluSerPheProGlyAsnLysSerAsnAsnAspGluAsnSerThrArgLysGlnSer 524

Qy       1590 ------AAGGAGAAGAAGAAGGTGGAAATGGCCGCCACT 1622
Db        525 AsnGlyAsnAspLysSerGlnValGlnLeuGlyGluThr 537

RESULT 13
US-10-051-902-30
; Sequence 30, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:
;  APPLICANT: Allen, Steve
;  APPLICANT:  Hitz, Bill
;  APPLICANT:  Kinney, Tony
;  APPLICANT:  Tingey, Scott
;  TITLE OF INVENTION: Plant Sugar Transport Proteins
;  FILE REFERENCE: BB-1163
;  CURRENT APPLICATION NUMBER: US/10/051,902
;  CURRENT FILING DATE:  2002-01-17
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
;  PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 30
;   LENGTH: 549
;   TYPE: PRT
;   ORGANISM: Beta vulgaris
US-10-051-902-30

Alignment Scores:
Pred. No.:              1.95e-86       Length:        549
```

Score:	1589.00	Matches:	326
Percent Similarity:	72.23%	Conservative:	59
Best Local Similarity:	61.16%	Mismatches:	126
Query Match:	43.97%	Indels:	22
DB:	13	Gaps:	5

US-10-051-902A-21 (1-2017) x US-10-051-902-30 (1-549)

```
Qy          78 TCCTCTGCACCACCGGAGATGGCTTCCGCCGCGCTG-----------CCGGAGGCCGTC 125
Db          10 SerAspProProProThrThrAlaSerLysValIleAlaAspPheAspProLeuLysLys 29

Qy         126 GCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATG 185
Db          30 ProProLysArgAsn---------LysPheAlaPheAlaCysAlaThrLeuAlaSerMet 46

Qy         186 ACCTCCATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAG 245
Db          47 ThrSerValLeuLeuGlyTyrAspIleGlyValMetSerGlyAlaIleIleTyrLeuLys 66

Qy         246 AAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTAC 305
Db          67 GluAspTrpHisIleSerAspThrGlnIleGlyValLeuValGlyIleLeuAsnIleTyr 86

Qy         306 TCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACC 365
Db          87 CysLeuPheGlySerPheAlaAlaGlyArgThrSerAspTrpIleGlyArgArgTyrThr 106

Qy         366 ATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAAC 425
Db         107 IleValLeuAlaGlyAlaIlePhePheValGlyAlaLeuLeuMetGlyPheAlaThrAsn 126

Qy         426 TACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATG 485
Db         127 TyrAlaPheLeuMetValGlyArgPheValThrGlyIleGlyValGlyTyrAlaLeuMet 146

Qy         486 ATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCG 545
Db         147 IleAlaProValTyrThrAlaGluValSerProAlaSerSerArgGlyPheLeuThrSer 166

Qy         546 TTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTC 605
Db         167 PheProGluValPheIleAsnAlaGlyIleLeuLeuGlyTyrIleSerAsnLeuAlaPhe 186

Qy         606 TCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCC 665
Db         187 SerSerLeuProThrHisLeuSerTrpArgPheMetLeuGlyIleGlyAlaIleProSer 206

Qy         666 GTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAG 725
Db         207 IlePheLeuAlaIleGlyValLeuAlaMetProGluSerProArgTrpLeuValMetGln 226

Qy         726 GGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCC 785
Db         227 GlyArgLeuGlyAspAlaLysLysValLeuAsnArgIleSerAspSerProGluGluAla 246

Qy         786 GCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGAC 845
Db         247 GlnLeuArgLeuSerGluIleLysGlnThrAlaGlyIleProAlaGluCysAspGluAsp 266

Qy         846 GTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATC 905
Db         267 IleTyrLysValGluLysThrLysIleLysSerGlyAsnAlaValTrpLysGluLeuPhe 286

Qy         906 CTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTC 965
Db         287 PheAsnProThrProAlaValArgArgAlaValIleAlaGlyIleGlyIleHisPhePhe 306

Qy         966 CAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCC 1025
Db         307 GlnGlnAlaSerGlyIleAspAlaValValLeuTyrSerProArgIlePheGlnSerAla 326

Qy        1026 GGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGG 1085
```

```
Db        327 GlyIleThrAsnAlaArgLysGlnLeuLeuAlaThrValAlaValGlyValValLysThr 346

Qy       1086 CTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTG 1145

Db        347 LeuPheIleLeuValAlaThrPheGlnLeuAspLysTyrGlyArgArgProLeuLeuLeu 366

Qy       1146 GGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTC 1205

Db        367 ThrSerValGlyGlyMetIleIleAlaIleLeuThrLeuAlaMetSerLeuThrValIle 386

Qy       1206 GGCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTC 1265

Db        387 ---AspHisSerHisHisLysIleThrTrpAlaIleAlaLeuCysIleThrMetValCys 405

Qy       1266 GCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAG 1325

Db        406 AlaValValAlaSerPheSerIleGlyLeuGlyProIleThrTrpValTyrSerSerGlu 425

Qy       1326 ATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTC 1385

Db        426 ValPheProLeuArgLeuArgAlaGlnGlyThrSerMetGlyValAlaValAsnArgVal 445

Qy       1386 ACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGC 1445

Db        446 ValSerGlyValIleSerIlePhePheLeuProLeuSerHisLysIleThrThrGlyGly 465

Qy       1446 AGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTC 1505

Db        466 AlaPhePheLeuPheGlyGlyIleAlaIleIleAlaTrpPhePhePheLeuThrPheLeu 485

Qy       1506 CCGGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGAC--------- 1556

Db        486 ProGluThrArgGlyArgThrLeuGluAsnMetHisGluLeuPheGluAspPheArgTrp 505

Qy       1557 ------------ACGGCCGCCGCCTCGGAATCAGACGAGCCAGCC--------------- 1589

Db        506 ArgGluSerPheProGlyAsnLysSerAsnAsnAspGluAsnSerThrArgLysGlnSer 525

Qy       1590 ------AAGGAGAAGAAGAAGGTGGAAATGGCCGCCACT 1622

Db        526 AsnGlyAsnAspLysSerGlnValGlnLeuGlyGluThr 538

RESULT 14
US-10-051-909-30
; Sequence 30, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
;  APPLICANT: Allen, Steve
;  APPLICANT:  Helentjaris, Tim
;  APPLICANT:  Hitz, Bill
;  APPLICANT:  Kinney, Tony
;  APPLICANT:  Tingey, Scott
;  TITLE OF INVENTION: Plant Sugar Transport Proteins
;  FILE REFERENCE: BB1163 US CIP
;  CURRENT APPLICATION NUMBER: US/10/051,909
;  CURRENT FILING DATE:  2002-01-17
;  PRIOR APPLICATION NUMBER: 60/083,044
;  PRIOR FILING DATE: April 24, 1998
;  NUMBER OF SEQ ID NOS: 38
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 30
;   LENGTH: 549
;   TYPE: PRT
;   ORGANISM: Beta vulgaris
US-10-051-909-30

Alignment Scores:
Pred. No.:              1.95e-86        Length:         549
Score:                  1589.00         Matches:        326
Percent Similarity:     72.23%          Conservative:   59
Best Local Similarity:  61.16%          Mismatches:     126
Query Match:            43.97%          Indels:         22

DB:                     13              Gaps:           5

US-10-051-902A-21 (1-2017) x US-10-051-909-30 (1-549)

Qy         78 TCCTCTGCACCACCGGAGATGGCTTCCGCCGCGCTG-----------CCGGAGGCCGTC 125

Db         10 SerAspProProProThrThrAlaSerLysValIleAlaAspPheAspProLeuLysLys 29

Qy        126 GCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATG 185

Db         30 ProProLysArgAsn---------LysPheAlaPheAlaCysAlaThrLeuAlaSerMet 46

Qy        186 ACCTCCATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAG 245

Db         47 ThrSerValLeuLeuGlyTyrAspIleGlyValMetSerGlyAlaIleIleTyrLeuLys 66

Qy        246 AAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTAC 305

Db         67 GluAspTrpHisIleSerAspThrGlnIleGlyValLeuValGlyIleLeuAsnIleTyr 86

Qy        306 TCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACC 365

Db         87 CysLeuPheGlySerPheAlaAlaGlyArgThrSerAspTrpIleGlyArgArgTyrThr 106

Qy        366 ATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAAC 425

Db        107 IleValLeuAlaGlyAlaIlePhePheValGlyAlaLeuLeuMetGlyPheAlaThrAsn 126

Qy        426 TACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATG 485

Db        127 TyrAlaPheLeuMetValGlyArgPheValThrGlyIleGlyValGlyTyrAlaLeuMet 146

Qy        486 ATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCG 545

Db        147 IleAlaProValTyrThrAlaGluValSerProAlaSerSerArgGlyPheLeuThrSer 166

Qy        546 TTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTC 605

Db        167 PheProGluValPheIleAsnAlaGlyIleLeuLeuGlyTyrIleSerAsnLeuAlaPhe 186

Qy        606 TCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCC 665

Db        187 SerSerLeuProThrHisLeuSerTrpArgPheMetLeuGlyIleGlyAlaIleProSer 206

Qy        666 GTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAG 725

Db        207 IlePheLeuAlaIleGlyValLeuAlaMetProGluSerProArgTrpLeuValMetGln 226

Qy        726 GGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCC 785

Db        227 GlyArgLeuGlyAspAlaLysLysValLeuAsnArgIleSerAspSerProGluGluAla 246

Qy        786 GCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGAC 845

Db        247 GlnLeuArgLeuSerGluIleLysGlnThrAlaGlyIleProAlaGluCysAspGluAsp 266

Qy        846 GTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATC 905

Db        267 IleTyrLysValGluLysThrLysIleLysSerGlyAsnAlaValTrpLysGluLeuPhe 286

Qy        906 CTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTC 965

Db        287 PheAsnProThrProAlaValArgArgAlaValIleAlaGlyIleGlyIleHisPhePhe 306

Qy        966 CAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCC 1025

Db        307 GlnGlnAlaSerGlyIleAspAlaValValLeuTyrSerProArgIlePheGlnSerAla 326

Qy       1026 GGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGG 1085

Db        327 GlyIleThrAsnAlaArgLysGlnLeuLeuAlaThrValAlaValGlyValValLysThr 346

Qy       1086 CTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTG 1145
```

```
Db        347 LeuPheIleLeuValAlaThrPheGlnLeuAspLysTyrGlyArgArgProLeuLeuLeu 366

Qy       1146 GGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTC 1205

Db        367 ThrSerValGlyGlyMetIleIleAlaIleLeuThrLeuAlaMetSerLeuThrValIle 386

Qy       1206 GGCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTC 1265

Db        387 ---AspHisSerHisHisLysIleThrTrpAlaIleAlaLeuCysIleThrMetValCys 405

Qy       1266 GCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAG 1325

Db        406 AlaValValAlaSerPheSerIleGlyLeuGlyProIleThrTrpValTyrSerSerGlu 425

Qy       1326 ATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTC 1385

Db        426 ValPheProLeuArgLeuArgAlaGlnGlyThrSerMetGlyValAlaValAsnArgVal 445

Qy       1386 ACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGC 1445

Db        446 ValSerGlyValIleSerIlePhePheLeuProLeuSerHisLysIleThrThrGlyGly 465

Qy       1446 AGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTC 1505

Db        466 AlaPhePheLeuPheGlyGlyIleAlaIleIleAlaTrpPhePhePheLeuThrPheLeu 485

Qy       1506 CCGGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGAC--------- 1556

Db        486 ProGluThrArgGlyArgThrLeuGluAsnMetHisGluLeuPheGluAspPheArgTrp 505

Qy       1557 ------------ACGGCCGCCGCCTCGGAATCAGACGAGCCAGCC--------------- 1589

Db        506 ArgGluSerPheProGlyAsnLysSerAsnAsnAspGluAsnSerThrArgLysGlnSer 525

Qy       1590 ------AAGGAGAAGAAGAAGGTGGAAATGGCCGCCACT 1622

Db        526 AsnGlyAsnAspLysSerGlnValGlnLeuGlyGluThr 538

RESULT 15
US-10-425-114-56035
; Sequence 56035, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 56035
;   LENGTH: 517
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: 701204606_FLI.pep
US-10-425-114-56035

Alignment Scores:
Pred. No.:                4.73e-86      Length:          517
Score:                    1582.50       Matches:         305
Percent Similarity:       77.53%        Conservative:    78
Best Local Similarity:    61.74%        Mismatches:      108
Query Match:              43.79%        Indels:          3
DB:                       12            Gaps:            2

US-10-051-902A-21 (1-2017) x US-10-425-114-56035 (1-517)
```

```
Qy        132 AAGAAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCC 191

Db         16 LysLysProArgArgAsnLysTyrAlaPheAlaCysAlaIleLeuAlaSerMetThrSer 35

Qy        192 ATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGAC 251

Db         36 IleLeuLeuGlyTyrAspIleGlyValMetSerGlyAlaAlaLeuTyrIleGlnArgAsp 55

Qy        252 TTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTC 311

Db         56 LeuLysValSerAspValGlnIleGluIleLeuAsnGlyIleIleAsnLeuTyrSerPro 75

Qy        312 ATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTG 371

Db         76 ValGlySerPheIleAlaGlyArgThrSerAspTrpIleGlyArgArgTyrThrIleVal 95

Qy        372 TTCGCCGCCGTCATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCC 431

Db         96 LeuAlaGlyAlaIlePhePheValGlyAlaIleLeuMetGlyPheSerProAsnTyrAla 115

Qy        432 ATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCG 491

Db        116 PheLeuMetPheGlyArgPhePheAlaGlyValGlyIleGlyPheAlaPheLeuIleAla 135

Qy        492 CCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCG 551

Db        136 ProValTyrThrSerGluIleSerProSerSerSerArgGlyPheLeuThrSerLeuPro 155

Qy        552 GAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGC 611

Db        156 GluValPheLeuAsnGlyGlyIleLeuIleGlyTyrIleSerAsnTyrGlyPheSerLys 175

Qy        612 TTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTG 671

Db        176 LeuAlaLeuArgLeuGlyTrpArgLeuMetLeuGlyValGlyAlaIleProSerIleLeu 195

Qy        672 CTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGC 731

Db        196 IleGlyValAlaValLeuAlaMetProGluSerProArgTrpLeuValAlaLysGlyArg 215

Qy        732 CTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAG 791

Db        216 LeuGlyGluAlaLysArgValLeuTyrLysIleSerGluSerGluGluGluAlaArgLeu 235

Qy        792 CGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTG 851

Db        236 ArgLeuAlaAspIleLysAspThrAlaGlyIleProGlnAspCysAspAspAspValVal 255

Qy        852 ACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCC 911

Db        256 LeuValSerLysGlnThrHisGlyHis------GlyValTrpArgGluLeuPheLeuHis 273

Qy        912 CCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCAT 971

Db        274 ProThrProAlaValArgHisIlePheIleAlaSerLeuGlyIleHisPhePheAlaGln 293

Qy        972 GCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTA 1031

Db        294 AlaThrGlyIleAspAlaValValLeuTyrSerProArgIlePheGluLysAlaGlyIle 313

Qy       1032 ACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTC 1091

Db        314 LysSerAspAsnTyrArgLeuLeuAlaThrValAlaValGlyPheValLysThrValSer 333

Qy       1092 ATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGC 1151

Db        334 IleLeuValAlaThrPhePheLeuAspArgAlaGlyArgArgValLeuLeuLeuCysSer 353

Qy       1152 ACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAG 1211

Db        354 ValSerGlyLeuIleLeuSerLeuLeuThrLeuGlyLeuSerLeuThrValVal---Asp 372
```

```
Qy       1212 CACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTAC 1271
              |||            :::   ||||||:::|||||||||||||||:::   |||:::|||
Db        373 HisSerGlnThrThrLeuAsnTrpAlaValGlyLeuSerIleAlaAlaValLeuSerTyr 392

Qy       1272 GTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTC 1331
              ||||||  ||||||||||||  ||||||||||||||||||||||||||||||||||||||
Db        393 ValAlaThrPheSerIleGlySerGlyProIleThrTrpValTyrSerSerGluIlePhe 412

Qy       1332 CCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGC 1391
              ||||||::::::||||||   |||   ::::::|||  |||   ||||||||||||||
Db        413 ProLeuArgLeuArgAlaGlnGlyValAlaIleGlyAlaAlaValAsnArgValThrSer 432

Qy       1392 GGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTC 1451
              |||||||||:::|||||||||||||||||  ||||||||||||||||||||:::|||
Db        433 GlyValIleAlaMetThrPheLeuSerLeuGlnLysAlaIleThrIleGlyGlyAlaPhe 452

Qy       1452 TTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAG 1511
              |||||:::::::|||:::||||||:::||||||:::|||  ||||||  |||||||||
Db        453 PheLeuPheAlaGlyValAlaAlaValAlaTrpIlePheHisTyrThrLeuLeuProGlu 472

Qy       1512 ACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCG 1571
              |||||||||:::|||||||||||:::  |||  ||||||:::
Db        473 ThrArgGlyLysThrLeuGluGluIleGluLysSerPheGlyAsnPheCysArgLysPro 492

Qy       1572 GAATCAGACGAGCCAGCCAAGGAGAAGAAGAAGGTGGAAATG 1613
              ::::::::::|||         :::      ::::::::::::
Db        493 LysAlaGluGluGlyLeuAspAspAsnValGluIleGlnLeu 506


Search completed: June 30, 2004, 19:09:22
Job time : 126 secs
```

GenCore version 5.1.6

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 30, 2004, 18:37:56 ; Search time 100 Seconds
(without alignments)
11397.966 Million cell updates/sec

Title: US-10-051-902A-21
Perfect score: 3614
Sequence: 1 cttacatgtaagctcgtgcc.....aaaaaaaaaaaaaaaaaaaa 2017

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 3172214

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10051902/runat_30062004_164722_20393/app_query.fasta_1.2183
-DB=A_Geneseq_29Jan04 -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10051902_@CGN_1_1_81_@runat_30062004_164722_20393 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2570	71.1	510	5	AAU97211	Aau97211 Rice Beta
2	2570	71.1	510	6	ABU08336	Abu08336 Rice suga
3	2015	55.8	539	5	AAU97213	Aau97213 Wheat sug
4	2015	55.8	539	6	ABU08338	Abu08338 Wheat sug
5	1877.5	52.0	529	5	AAU97214	Aau97214 Wheat sug
6	1877.5	52.0	529	6	ABU08339	Abu08339 Wheat sug
7	1872.5	51.8	513	5	AAU97210	Aau97210 Corn Beta
8	1872.5	51.8	513	6	ABU08335	Abu08335 Corn suga
9	1660	45.9	523	5	AAU97212	Aau97212 Soybean B
10	1660	45.9	523	6	ABU08337	Abu08337 Soybean s
11	1513	41.9	513	5	ABB09681	Abb09681 Amino aci
12	1385.5	38.3	491	3	AAG32072	Aag32072 Arabidops
13	1385.5	38.3	508	3	AAG32071	Aag32071 Arabidops
14	1303.5	36.1	466	3	AAG32073	Aag32073 Arabidops
15	827.5	22.9	333	6	ADA48320	Ada48320 Rice prot
16	682	18.9	167	5	AAU97209	Aau97209 Portion o
17	682	18.9	167	6	ABU08334	Abu08334 Corn suga
18	679.5	18.8	487	6	ABG73334	Abg73334 Consensus
19	672	18.6	488	6	ABP98504	Abp98504 PFAM cons
20	672	18.6	488	6	AAE35306	Aae35306 Human sug
21	672	18.6	488	7	ADD22918	Add22918 Human sug
22	658.5	18.2	629	4	AAB66935	Aab66935 Human GLU
23	658.5	18.2	629	4	AAB66940	Aab66940 GLUTX2 co
24	653.5	18.1	648	5	ABG61548	Abg61548 Human tra
25	649	18.0	446	6	ABU43281	Abu43281 Protein e
26	639.5	17.7	618	4	AAB66936	Aab66936 Rat GLUTX
27	629.5	17.4	582	3	AAG29528	Aag29528 Arabidops
28	628.5	17.4	555	3	AAG29529	Aag29529 Arabidops
29	594	16.4	517	5	ABP66239	Abp66239 Bifidobac
30	591	16.4	478	3	AAG15416	Aag15416 Arabidops
31	591	16.4	493	3	AAG15415	Aag15415 Arabidops
32	591	16.4	546	3	AAG15414	Aag15414 Arabidops
33	584.5	16.2	450	6	ABU33853	Abu33853 Protein e
34	582	16.1	491	4	AAG89949	Aag89949 C glutami
35	580	16.0	491	3	AAB12594	Aab12594 Brevibact
36	579.5	16.0	472	5	ABP52164	Abp52164 E. coli a
37	579.5	16.0	472	7	ADD37502	Add37502 E. coli a
38	576	15.9	584	5	ABB05603	Abb05603 Yeast ino
39	572.5	15.8	464	5	ABP52163	Abp52163 E. coli g
40	572.5	15.8	464	6	ABU14994	Abu14994 Protein e
41	572.5	15.8	464	7	ADD37501	Add37501 E. coli g
42	571.5	15.8	465	6	ABU28442	Abu28442 Protein e
43	571	15.8	464	6	ABU48384	Abu48384 Protein e
44	571	15.8	464	6	ABU46992	Abu46992 Protein e
45	569	15.7	453	6	ABU31194	Abu31194 Protein e

ALIGNMENTS

RESULT 1
AAU97211
ID AAU97211 standard; protein; 510 AA.
XX
AC AAU97211;
XX
DT 27-AUG-2002 (first entry)
XX
DE Rice Beta vulgaris-like sugar transport protein.
XX
KW Rice; Beta vulgaris-like sugar transport protein; carbohydrate transport;
KW grain filling; annual field crop; plant.
XX
OS Oryza sativa.
XX
FH Key Location/Qualifiers
FT Misc-difference 102
FT /label= Unknown
XX
PN US6383776-B1.
XX
PD 07-MAY-2002.
XX
PF 14-APR-1999; 99US-00291922.
XX
PR 24-APR-1998; 98US-0083044P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX
DR WPI; 2002-453364/48.
DR N-PSDB; ABK51972.
XX

PT New nucleic acid encoding plant sugar-transport proteins, useful for
PT preparing transgenic plants with altered carbohydrate distribution.
XX
PS Example 4; Fig 2; 54pp; English.
XX
CC The present invention relates to the isolation of plant polynucleotide
CC sequences encoding an Arabidopsis thaliana-like sugar transport protein
CC or Beta vulgaris-like sugar transport protein. The polynucleotide
CC sequences are useful for altering the level of sugar transport proteins
CC in plants, i.e. for control of carbohydrate transport and distribution in
CC plant cells, e.g. during grain filling of annual field crops (e.g. corn,
CC rice, soybeans, and wheat), and, for studying carbohydrate flows and
CC sugar transport. The polynucleotide sequences can also be used to isolate
CC cDNA sequences and genes that encode homologues of the new proteins. The
CC present sequence represents a rice Beta vulgaris-like sugar transport
CC protein
XX
SQ Sequence 510 AA;

Alignment Scores:

Pred. No.:	9.06e-186	Length:	510
Score:	2570.00	Matches:	510
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	71.11%	Indels:	0
DB:	5	Gaps:	0

US-10-051-902A-21 (1-2017) x AAU97211 (1-510)

```
Qy         96 ATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTC 155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetAlaSerAlaAlaLeuProGluAlaValAlaProLysLysLysGlyAsnValArgPhe 20

Qy        156 GCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGG 215
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AlaPheAlaCysAlaIleLeuAlaSerMetThrSerIleLeuLeuGlyTyrAspIleGly 40

Qy        216 GTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTG 275
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ValMetSerGlyAlaSerLeuTyrIleLysLysAspPheAsnIleSerAspGlyLysVal 60

Qy        276 GAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGG 335
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GluValLeuMetGlyIleLeuAsnLeuTyrSerLeuIleGlySerPheAlaAlaGlyArg 80

Qy        336 ACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCG 395
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ThrSerAspTrpIleGlyArgArgTyrThrIleValPheAlaAlaValIlePhePheAla 100

Qy        396 GGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTG 455
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 Gly***PheLeuMetGlyPheAlaValAsnTyrAlaMetLeuMetPheGlyArgPheVal 120

Qy        456 GCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCG 515
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AlaGlyIleGlyValGlyTyrAlaLeuMetIleAlaProValTyrThrAlaGluValSer 140

Qy        516 CCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATC 575
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 ProAlaSerAlaArgGlyPheLeuThrSerPheProGluValPheIleAsnPheGlyIle 160

Qy        576 CTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGC 635
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 LeuLeuGlyTyrValSerAsnTyrAlaPheSerArgLeuProLeuAsnLeuGlyTrpArg 180

Qy        636 ATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATG 695
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IleMetLeuGlyIleGlyAlaAlaProSerValLeuLeuAlaLeuMetValLeuGlyMet 200

Qy        696 CCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTG 755
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ProGluSerProArgTrpLeuValMetLysGlyArgLeuAlaAspAlaLysValValLeu 220

Qy        756 GAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCC 815
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 GluLysThrSerAspThrAlaGluGluAlaAlaGluArgLeuAlaAspIleLysAlaAla 240

Qy        816 GCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGGAGCGGA 875
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AlaGlyIleProGluGluLeuAspGlyAspValValThrValProLysArgGlySerGly 260

Qy        876 AACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATC 935
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 AsnGluLysArgValTrpLysGluLeuIleLeuSerProThrProAlaMetArgArgIle 280

Qy        936 CTGCTGTCCGGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTC 995
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 LeuLeuSerGlyIleGlyIleHisPhePheGlnHisAlaLeuGlyIleHisSerValVal 300

Qy        996 TTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGC 1055
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 PheTyrSerProLeuValPheLysSerProGlyLeuThrAsnAspLysHisPheLeuGly 320

Qy       1056 ACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATC 1115
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ThrThrTrpProPheGlyValThrLysArgLeuPheIleLeuLeuAlaThrPhePheIle 340

Qy       1116 GACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTC 1175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 AspGlyValGlyArgArgProLeuLeuLeuGlySerThrGlyGlyIleIleLeuSerLeu 360

Qy       1176 ATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTGG 1235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 IleGlyLeuGlyAlaGlyLeuThrValValGlyGlnHisProAspAlaLysIleProTrp 380

Qy       1236 GCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTT 1295
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 AlaIleGlyLeuSerIleAlaSerThrLeuAlaTyrValAlaPhePheSerIleGlyLeu 400

Qy       1296 GGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGC 1355
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 GlyProIleThrTrpValTyrSerSerGluIlePheProLeuGlnValArgAlaLeuGly 420

Qy       1356 TGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTG 1415
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CysSerLeuGlyValAlaAlaAsnArgValThrSerGlyValIleSerMetThrPheLeu 440

Qy       1416 TCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCG 1475
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 SerLeuSerLysAlaIleThrIleGlyGlySerPhePheLeuTyrSerGlyIleAlaAla 460

Qy       1476 CTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAG 1535
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 LeuAlaTrpValPhePheTyrThrTyrLeuProGluThrArgGlyArgThrLeuGluGlu 480

Qy       1536 ATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAGCCAAGGAG 1595
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 MetSerLysLeuPheGlyAspThrAlaAlaAlaSerGluSerAspGluProAlaLysGlu 500

Qy       1596 AAGAAGAAGGTGGAAATGGCCGCCACTAAC 1625
              |||||||||||||||||||||||||||||
Db        501 LysLysLysValGluMetAlaAlaThrAsn 510
```

RESULT 2
ABU08336
ID ABU08336 standard; protein; 510 AA.
XX
AC ABU08336;
XX
DT 29-MAY-2003 (first entry)
XX
DE Rice sugar transport protein #3.
XX
KW Beta vulgaris-like sugar transport protein; corn; rice; wheat;
KW plant sugar transport protein; carbohydrate transport; soybean;
KW carbohydrate distribution; plant.

XX
OS Oryza sativa.
XX
PN US2002178468-A1.
XX
PD 28-NOV-2002.
XX
PF 17-JAN-2002; 2002US-00051902.
XX
PR 24-APR-1998; 98US-0083044P.
PR 14-APR-1999; 99US-00291922.
XX
PA (ALLE/) ALLEN S M.
PA (HITZ/) HITZ W D.
PA (KINN/) KINNEY A J.
PA (TING/) TINGEY S V.
XX
PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX
DR WPI; 2003-340957/32.
DR N-PSDB; ABX93208.
XX
PT Novel plant sugar transport proteins and nucleic acid encoding the
PT protein useful for producing transgenic plants having altered levels of
PT sugar transport protein.
XX
PS Claim 10; Fig 2; 56pp; English.
XX
CC The present invention relates to the isolation of Arabidopsis thaliana-
CC like or Beta vulgaris-like sugar transport proteins, and the
CC polynucleotide sequences encoding them. The plant sugar transport
CC proteins of the invention have been isolated from corn, rice, soybean,
CC and wheat. The polypeptides of the invention may be used for altering the
CC level of expression of a sugar transport protein in a host cell, by
CC transforming a host cell with a chimeric construct encoding all, or a
CC portion of the sugar transport protein, in sense or antisense
CC orientation. Particularly, the polypeptides may provide a means to
CC control carbohydrate transport and distribution in plants. ABU08334-
CC ABU08339 represent Beta vulgaris-like sugar transport proteins
XX
SQ Sequence 510 AA;

Alignment Scores:

Pred. No.:	9.06e-186	Length:	510
Score:	2570.00	Matches:	510
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	71.11%	Indels:	0
DB:	6	Gaps:	0

US-10-051-902A-21 (1-2017) x ABU08336 (1-510)

```
Qy          96 ATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTC 155
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 MetAlaSerAlaAlaLeuProGluAlaValAlaProLysLysLysGlyAsnValArgPhe 20

Qy         156 GCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGG 215
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          21 AlaPheAlaCysAlaIleLeuAlaSerMetThrSerIleLeuLeuGlyTyrAspIleGly 40

Qy         216 GTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTG 275
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          41 ValMetSerGlyAlaSerLeuTyrIleLysLysAspPheAsnIleSerAspGlyLysVal 60

Qy         276 GAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGGCGG 335
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 GluValLeuMetGlyIleLeuAsnLeuTyrSerLeuIleGlySerPheAlaAlaGlyArg 80

Qy         336 ACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCG 395
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          81 ThrSerAspTrpIleGlyArgArgTyrThrIleValPheAlaAlaValIlePhePheAla 100

Qy         396 GGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTG 455
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         101 Gly***PheLeuMetGlyPheAlaValAsnTyrAlaMetLeuMetPheGlyArgPheVal 120

Qy         456 GCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCG 515
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 AlaGlyIleGlyValGlyTyrAlaLeuMetIleAlaProValTyrThrAlaGluValSer 140

Qy         516 CCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATC 575
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         141 ProAlaSerAlaArgGlyPheLeuThrSerPheProGluValPheIleAsnPheGlyIle 160

Qy         576 CTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGC 635
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         161 LeuLeuGlyTyrValSerAsnTyrAlaPheSerArgLeuProLeuAsnLeuGlyTrpArg 180

Qy         636 ATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATG 695
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 IleMetLeuGlyIleGlyAlaAlaProSerValLeuLeuAlaLeuMetValLeuGlyMet 200

Qy         696 CCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTG 755
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         201 ProGluSerProArgTrpLeuValMetLysGlyArgLeuAlaAspAlaLysValValLeu 220

Qy         756 GAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCC 815
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         221 GluLysThrSerAspThrAlaGluGluAlaAlaGluArgLeuAlaAspIleLysAlaAla 240

Qy         816 GCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGA 875
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 AlaGlyIleProGluGluLeuAspGlyAspValValThrValProLysArgGlySerGly 260

Qy         876 AACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATC 935
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         261 AsnGluLysArgValTrpLysGluLeuIleLeuSerProThrProAlaMetArgArgIle 280

Qy         936 CTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTC 995
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         281 LeuLeuSerGlyIleGlyIleHisPhePheGlnHisAlaLeuGlyIleHisSerValVal 300

Qy         996 TTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGC 1055
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 PheTyrSerProLeuValPheLysSerProGlyLeuThrAsnAspLysHisPheLeuGly 320

Qy        1056 ACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATC 1115
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         321 ThrThrTrpProPheGlyValThrLysArgLeuPheIleLeuLeuAlaThrPhePheIle 340

Qy        1116 GACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTC 1175
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         341 AspGlyValGlyArgArgProLeuLeuLeuGlySerThrGlyGlyIleIleLeuSerLeu 360

Qy        1176 ATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTGG 1235
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 IleGlyLeuGlyAlaGlyLeuThrValValGlyGlnHisProAspAlaLysIleProTrp 380

Qy        1236 GCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTT 1295
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         381 AlaIleGlyLeuSerIleAlaSerThrLeuAlaTyrValAlaPhePheSerIleGlyLeu 400

Qy        1296 GGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGC 1355
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         401 GlyProIleThrTrpValTyrSerSerGluIlePheProLeuGlnValArgAlaLeuGly 420

Qy        1356 TGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTG 1415
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 CysSerLeuGlyValAlaAlaAsnArgValThrSerGlyValIleSerMetThrPheLeu 440

Qy        1416 TCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCG 1475
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         441 SerLeuSerLysAlaIleThrIleGlyGlySerPhePheLeuTyrSerGlyIleAlaAla 460

Qy        1476 CTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAG 1535
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        461 LeuAlaTrpValPhePheTyrThrTyrLeuProGluThrArgGlyArgThrLeuGluGlu 480

Qy       1536 ATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAGCCAAGGAG 1595
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 MetSerLysLeuPheGlyAspThrAlaAlaAlaSerGluSerAspGluProAlaLysGlu 500

Qy       1596 AAGAAGAAGGTGGAAATGGCCGCCACTAAC 1625
              |||||||||||||||||||||||||||||
Db        501 LysLysLysValGluMetAlaAlaThrAsn 510
```

RESULT 3
AAU97213

ID AAU97213 standard; protein; 539 AA.
XX
AC AAU97213;
XX
DT 27-AUG-2002 (first entry)
XX
DE Wheat sugar transport protein encoded by wlk8.pk0001.a11.
XX
KW Wheat; Beta vulgaris-like sugar transport protein;
KW carbohydrate transport; grain filling; annual field crop; plant.
XX
OS Triticum aestivum.
XX
PN US6383776-B1.
XX
PD 07-MAY-2002.
XX
PF 14-APR-1999; 99US-00291922.
XX
PR 24-APR-1998; 98US-0083044P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX
DR WPI; 2002-453364/48.
DR N-PSDB; ABK51974.
XX
PT New nucleic acid encoding plant sugar-transport proteins, useful for
PT preparing transgenic plants with altered carbohydrate distribution.
XX
PS Example 4; Fig 2; 54pp; English.
XX
CC The present invention relates to the isolation of plant polynucleotide
CC sequences encoding an Arabidopsis thaliana-like sugar transport protein
CC or Beta vulgaris-like sugar transport protein. The polynucleotide
CC sequences are useful for altering the level of sugar transport proteins
CC in plants, i.e. for control of carbohydrate transport and distribution in
CC plant cells, e.g. during grain filling of annual field crops (e.g. corn,
CC rice, soybeans, and wheat), and, for studying carbohydrate flows and
CC sugar transport. The polynucleotide sequences can also be used to isolate
CC cDNA sequences and genes that encode homologues of the new proteins. The
CC present sequence represents a wheat Beta vulgaris-like sugar transport
CC protein
XX
SQ Sequence 539 AA;

Alignment Scores:

Pred. No.:	9.75e-144	Length:	539
Score:	2015.00	Matches:	408
Percent Similarity:	84.49%	Conservative:	44
Best Local Similarity:	76.26%	Mismatches:	73
Query Match:	55.76%	Indels:	10
DB:	5	Gaps:	4

US-10-051-902A-21 (1-2017) x AAU97213 (1-539)

```
Qy         48 ACTGTACACGGCCCAGAGCGA-----------GCCTCCTCCTCCTCTGCACCACCGGAG 95
                 |||      ||||||   |||           |||:::|||      |||
Db          6 ThrGlnGlyGlyProArgArgHisAsnProGlnAlaAlaSerArgGlyAlaSerSerThr 25

Qy         96 ATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTC 155
                |||      ||||||||||||   ||||||   ||||||||||||||||||||||||||||
Db         26 MetAspArgAlaAlaLeuProAlaAlaValGluProLysLysLysGlyAsnValArgPhe 45

Qy        156 GCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGG 215
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         46 AlaPheAlaCysAlaIleLeuAlaSerMetThrSerIleLeuLeuGlyTyrAspIleGly 65

Qy        216 GTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTG 275
                ||||||||||||||||||||||||||:::||||||      |||:::|||   ::::::
Db         66 ValMetSerGlyAlaSerLeuTyrIleGlnLysAspLeuLysIleAsnAspThrGlnLeu 85

Qy        276 GAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGG 335
                |||||||||||||||||||||||:::|||||||||||||||||||||||||||||||||
Db         86 GluValLeuMetGlyIleLeuAsnValTyrSerLeuIleGlySerPheAlaAlaGlyArg 105

Qy        336 ACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCG 395
                |||||||||||||||||||||||||:::|||||||||||||||||||||||||||||||
Db        106 ThrSerAspTrpIleGlyArgArgPheThrIleValPheAlaAlaValIlePhePheAla 125

Qy        396 GGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTG 455
                |||      :::|||||||||:::|||||||||||||||||||||||||||||||||||
Db        126 GlyAlaLeuIleMetGlyPheSerValAsnTyrAlaMetLeuMetPheGlyArgPheVal 145

Qy        456 GCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCG 515
                |||||||||||||||||||||||||||||||||||||||||||||||   |||   |||||||||
Db        146 AlaGlyIleGlyValGlyTyrAlaLeuMetIleAlaProValAsnThrGlyGluValSer 165

Qy        516 CCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATC 575
                |||||||||||||||||   ||||||||||||||||||||||||||||||||||||||||
Db        166 ProAlaSerAlaArgGlyValLeuThrSerPheProGluValPheIleAsnPheGlyIle 185

Qy        576 CTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGC 635
                |||||||||||||||||||:::||||||:::||||||   |||   ||||||||||||
Db        186 LeuLeuGlyTyrValSerAsnPheAlaPheAlaArgLeuSerLeuArgLeuGlyTrpArg 205

Qy        636 ATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATG 695
                |||||||||||||||||||   ||||||||||||||||||   ||||||||||||||||
Db        206 IleMetLeuGlyIleGlyAlaValProSerValLeuLeuAlaPheMetValLeuGlyMet 225

Qy        696 CCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTG 755
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        226 ProGluSerProArgTrpLeuValMetLysGlyArgLeuAlaAspAlaLysValValLeu 245

Qy        756 GAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCC 815
                  ||||||||||||||   ||||||||||||||||||:::|||||||||||   |||
Db        246 AlaLysThrSerAspThrProGluGluAlaAlaGluArgIleAlaAspIleLysThrAla 265

Qy        816 GCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAG---AGAGGGAGC 872
                |||||||||||      |||||||||||||||||   |||||||||   :::||||||
Db        266 AlaGlyIleProLeuGlyLeuAspGlyAspValValProValProLysAsnLysGlySer 285

Qy        873 GGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGC 932
                  ||||||||||||   |||:::|||||||||||||||||||   |||||||||
Db        286 SerGluGluLysArgValLeuLysAspLeuIleLeuSerProThrIleAlaMetArgHis 305

Qy        933 ATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTC 992
                ||||||::::::|||||||||||||||||||||||   :::   ||||||   :::|||
Db        306 IleLeuIleAlaGlyIleGlyIleHisPhePheGlnGlnSerSerGlyIleAspAlaVal 325

Qy        993 GTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTG 1052
                |||   |||||||||||||||||||||||   |||:::|||   |||
Db        326 ValLeuTyrSerProLeuValPheLysSerAlaGlyIleThrGlyAspSerArgLeuArg 345

Qy       1053 GGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTC 1112
                |||||||||      |||   |||      :::|||||||||:::|||||||||
Db        346 GlyThrThrValAlaValGlyAlaThrAsnThrValPheIleLeuValAlaThrPheLeu 365

Qy       1113 ATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGGATAATCCTCTCC 1172
                :::|||   :::   ||||||||||||:::|||   ||||||||||:::::::::|||
Db        366 LeuAspArgIleArgArgArgProLeuValLeuThrSerThrGlyGlyMetLeuValSer 385

Qy       1173 CTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCT 1232
```

```
Db        386 LeuValGlyLeuAlaThrGlyLeuThrValIleSerArgHisProAspGluLysIleThr 405

Qy       1233 TGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGC 1292
Db        406 TrpAlaIleValLeuCysIlePheCysIleMetAlaTyrValAlaPhePheSerIleGly 425

Qy       1293 CTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTG 1352
Db        426 LeuGlyProIleThrTrpValTyrSerSerGluIlePheProLeuHisValArgAlaLeu 445

Qy       1353 GGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTC 1412
Db        446 GlyCysSerLeuGlyValAlaValAsnArgLeuThrSerGlyValIleSerMetThrPhe 465

Qy       1413 CTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCC 1472
Db        466 IleSerLeuSerLysAlaMetThrIleGlyGlyAlaPhePheLeuPheAlaGlyIleAla 485

Qy       1473 GCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAG 1532
Db        486 SerPheAlaTrpValPhePhePheAlaTyrLeuProGluThrArgGlyArgThrLeuGlu 505

Qy       1533 GAGATGAGCAAGCTGTTCGGCGACACGGCC------------GCCGCCTCGGAATCAGAC 1580
Db        506 AspMetSerSerLeuPheGlyAsnThrAlaThrHisLysGlnGlyAlaAlaGluAlaAsp 525

Qy       1581 GAGCCAGCCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCACTAAC 1625
Db        526 AspAspAlaGlyGlu---LysLysValGluMetAlaAlaThrAsn 539

RESULT 4
ABU08338
ID   ABU08338 standard; protein; 539 AA.
XX
AC   ABU08338;
XX
DT   29-MAY-2003  (first entry)
XX
DE   Wheat sugar transport protein #4.
XX
KW   Beta vulgaris-like sugar transport protein; corn; rice; wheat;
KW   plant sugar transport protein; carbohydrate transport; soybean;
KW   carbohydrate distribution; plant.
XX
OS   Triticum aestivum.
XX
PN   US2002178468-A1.
XX
PD   28-NOV-2002.
XX
PF   17-JAN-2002; 2002US-00051902.
XX
PR   24-APR-1998;   98US-0083044P.
PR   14-APR-1999;   99US-00291922.
XX
PA   (ALLE/) ALLEN S M.
PA   (HITZ/) HITZ W D.
PA   (KINN/) KINNEY A J.
PA   (TING/) TINGEY S V.
XX
PI   Allen SM,  Hitz WD,  Kinney AJ,  Tingey SV;
XX
DR   WPI; 2003-340957/32.
DR   N-PSDB; ABX93210.
XX
PT   Novel plant sugar transport proteins and nucleic acid encoding the
PT   protein useful for producing transgenic plants having altered levels of
PT   sugar transport protein.
XX
PS   Claim 10; Fig 2; 56pp; English.
XX
CC   The present invention relates to the isolation of Arabidopsis thaliana-
CC   like or Beta vulgaris-like sugar transport proteins, and the
CC   polynucleotide sequences encoding them. The plant sugar transport
CC   proteins of the invention have been isolated from corn, rice, soybean,
CC   and wheat. The polypeptides of the invention may be used for altering the
CC   level of expression of a sugar transport protein in a host cell, by
CC   transforming a host cell with a chimeric construct encoding all, or a
CC   portion of the sugar transport protein, in sense or antisense
CC   orientation. Particularly, the polypeptides may provide a means to
CC   control carbohydrate transport and distribution in plants. ABU08334-
CC   ABU08339 represent Beta vulgaris-like sugar transport proteins
XX
SQ   Sequence 539 AA;

Alignment Scores:
Pred. No.:                9.75e-144      Length:          539
Score:                    2015.00        Matches:         408
Percent Similarity:       84.49%         Conservative:    44
Best Local Similarity:    76.26%         Mismatches:      73
Query Match:              55.76%         Indels:          10
DB:                       6              Gaps:            4

US-10-051-902A-21 (1-2017) x ABU08338 (1-539)

Qy         48 ACTGTACACGGCCCAGAGCGA-----------GCCTCCTCCTCCTCTGCACCACCGGAG 95
Db          6 ThrGlnGlyGlyProArgArgHisAsnProGlnAlaAlaSerArgGlyAlaSerSerThr 25

Qy         96 ATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTC 155
Db         26 MetAspArgAlaAlaLeuProAlaAlaValGluProLysLysLysGlyAsnValArgPhe 45

Qy        156 GCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGG 215
Db         46 AlaPheAlaCysAlaIleLeuAlaSerMetThrSerIleLeuLeuGlyTyrAspIleGly 65

Qy        216 GTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTG 275
Db         66 ValMetSerGlyAlaSerLeuTyrIleGlnLysAspLeuLysIleAsnAspThrGlnLeu 85

Qy        276 GAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGG 335
Db         86 GluValLeuMetGlyIleLeuAsnValTyrSerLeuIleGlySerPheAlaAlaGlyArg 105

Qy        336 ACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCG 395
Db        106 ThrSerAspTrpIleGlyArgArgPheThrIleValPheAlaAlaValIlePhePheAla 125

Qy        396 GGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTG 455
Db        126 GlyAlaLeuIleMetGlyPheSerValAsnTyrAlaMetLeuMetPheGlyArgPheVal 145

Qy        456 GCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCG 515
Db        146 AlaGlyIleGlyValGlyTyrAlaLeuMetIleAlaProValAsnThrGlyGluValSer 165

Qy        516 CCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATC 575
Db        166 ProAlaSerAlaArgGlyValLeuThrSerPheProGluValPheIleAsnPheGlyIle 185

Qy        576 CTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGC 635
Db        186 LeuLeuGlyTyrValSerAsnPheAlaPheAlaArgLeuSerLeuArgLeuGlyTrpArg 205

Qy        636 ATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATG 695
Db        206 IleMetLeuGlyIleGlyAlaValProSerValLeuLeuAlaPheMetValLeuGlyMet 225

Qy        696 CCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTG 755
Db        226 ProGluSerProArgTrpLeuValMetLysGlyArgLeuAlaAspAlaLysValValLeu 245

Qy        756 GAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCC 815
```

```
Db        246 AlaLysThrSerAspThrProGluGluAlaAlaGluArgIleAlaAspIleLysThrAla 265

Qy        816 GCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAG---AGAGGGAGC 872
              ||||||||||||      ||||||||||||||||||    |||||||||   :::||||||
Db        266 AlaGlyIleProLeuGlyLeuAspGlyAspValValProValProLysAsnLysGlySer 285

Qy        873 GGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGC 932
              ||||||||||||   |||:::||||||||||||||||||   |||||||||
Db        286 SerGluGluLysArgValLeuLysAspLeuIleLeuSerProThrIleAlaMetArgHis 305

Qy        933 ATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTC 992
              ||||||::::::||||||||||||||||||||||||   :::   ||||||   :::|||
Db        306 IleLeuIleAlaGlyIleGlyIleHisPhePheGlnGlnSerSerGlyIleAspAlaVal 325

Qy        993 GTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTG 1052
              |||   ||||||||||||||||||||||||   |||:::|||   |||
Db        326 ValLeuTyrSerProLeuValPheLysSerAlaGlyIleThrGlyAspSerArgLeuArg 345

Qy       1053 GGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTC 1112
              |||||||||         |||   |||      :::||||||||:::|||||||||
Db        346 GlyThrThrValAlaValGlyAlaThrAsnThrValPheIleLeuValAlaThrPheLeu 365

Qy       1113 ATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCC 1172
              :::|||   :::   ||||||||||||:::|||   ||||||||||:::::::::|||
Db        366 LeuAspArgIleArgArgArgProLeuValLeuThrSerThrGlyGlyMetLeuValSer 385

Qy       1173 CTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCT 1232
              |||:::||||||      ||||||||||:::   :::|||||||||   ||||||
Db        386 LeuValGlyLeuAlaThrGlyLeuThrValIleSerArgHisProAspGluLysIleThr 405

Qy       1233 TGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGC 1292
              |||||||||   |||   |||         :::|||||||||||||||||||||||||||
Db        406 TrpAlaIleValLeuCysIlePheCysIleMetAlaTyrValAlaPhePheSerIleGly 425

Qy       1293 CTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTG 1352
              ||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db        426 LeuGlyProIleThrTrpValTyrSerSerGluIlePheProLeuHisValArgAlaLeu 445

Qy       1353 GGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTC 1412
              ||||||||||||||||||   ||||||:::||||||||||||||||||||||||||||
Db        446 GlyCysSerLeuGlyValAlaValAsnArgLeuThrSerGlyValIleSerMetThrPhe 465

Qy       1413 CTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCC 1472
              :::|||||||||||||||:::|||||||||||:::|||||||||::::::|||||||||
Db        466 IleSerLeuSerLysAlaMetThrIleGlyGlyAlaPhePheLeuPheAlaGlyIleAla 485

Qy       1473 GCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAG 1532
              :::   |||||||||||||||:::   ||||||||||||||||||||||||||||||||
Db        486 SerPheAlaTrpValPhePhePheAlaTyrLeuProGluThrArgGlyArgThrLeuGlu 505

Qy       1533 GAGATGAGCAAGCTGTTCGGCGACACGGCC-----------GCCGCCTCGGAATCAGAC 1580
              :::||||||   |||||||||:::||||||                |||:::|||:::|||
Db        506 AspMetSerSerLeuPheGlyAsnThrAlaThrHisLysGlnGlyAlaAlaGluAlaAsp 525

Qy       1581 GAGCCAGCCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCACTAAC 1625
              :::   |||   |||   |||||||||||||||||||||||||
Db        526 AspAspAlaGlyGlu---LysLysValGluMetAlaAlaThrAsn 539

RESULT 5
AAU97214
ID   AAU97214 standard; protein; 529 AA.
XX
AC   AAU97214;
XX
DT   27-AUG-2002  (first entry)
XX
DE   Wheat sugar transport protein encoded by wlm1.pk0012.h1.
XX
KW   Wheat; Beta vulgaris-like sugar transport protein;
KW   carbohydrate transport; grain filling; annual field crop; plant.
XX
OS   Triticum aestivum.
XX
PN   US6383776-B1.
XX
PD   07-MAY-2002.
XX
PF   14-APR-1999;   99US-00291922.
XX
PR   24-APR-1998;   98US-0083044P.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Allen SM,  Hitz WD,  Kinney AJ,  Tingey SV;
XX
DR   WPI; 2002-453364/48.
DR   N-PSDB; ABK51975.
XX
PT   New nucleic acid encoding plant sugar-transport proteins, useful for
PT   preparing transgenic plants with altered carbohydrate distribution.
XX
PS   Example 4; Col 77-80; 54pp; English.
XX
CC   The present invention relates to the isolation of plant polynucleotide
CC   sequences encoding an Arabidopsis thaliana-like sugar transport protein
CC   or Beta vulgaris-like sugar transport protein. The polynucleotide
CC   sequences are useful for altering the level of sugar transport proteins
CC   in plants, i.e. for control of carbohydrate transport and distribution in
CC   plant cells, e.g. during grain filling of annual field crops (e.g. corn,
CC   rice, soybeans, and wheat), and, for studying carbohydrate flows and
CC   sugar transport. The polynucleotide sequences can also be used to isolate
CC   cDNA sequences and genes that encode homologues of the new proteins. The
CC   present sequence represents a wheat Beta vulgaris-like sugar transport
CC   protein
XX
SQ   Sequence 529 AA;

Alignment Scores:
Pred. No.:                2.5e-133        Length:          529
Score:                    1877.50         Matches:         369
Percent Similarity:       81.38%          Conservative:    55
Best Local Similarity:    70.83%          Mismatches:      92
Query Match:              51.95%          Indels:          5
DB:                       5               Gaps:            3

US-10-051-902A-21 (1-2017) x AAU97214 (1-529)

Qy         60 CCAGAGCGA------GCCTCCTCCTCCTCTGCACCACCGGAGATGGCTTCCGCCGCGCTG 113
              |||||||||      |||                              ||||||||||||||||||
Db          5 ProGluArgLysGlyAlaGluAspLysGluGluGlySerArgMetAlaSerAlaAlaLeu 24

Qy        114 CCGGAG------GCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGC 167
              ||||||      ||||||   |||:::   |||||||||   ::::::||||||   |||
Db         25 ProGluProGlyAlaValHisProArgAsnLysGlyAsnPheLysTyrAlaPheThrCys 44

Qy        168 GCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGGGGTGATGAGCGGG 227
              |||:::   |||||||||   :::|||:::|||||||||||||:::|||||||||||||||
Db         45 AlaLeuCysAlaSerMetAlaThrIleValLeuGlyTyrAspValGlyValMetSerGly 64

Qy        228 GCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATG 287
              ||||||||||||||||||:::|||      |||:::|||   ::::::|||::::::|||
Db         65 AlaSerLeuTyrIleLysArgAspLeuGlnIleThrAspValGlnLeuGluIleMetMet 84

Qy        288 GGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGG 347
              |||||||||:::::|||:::||||||||||||||||      ||||||||||||||
Db         85 GlyIleLeuSerValTyrAlaLeuIleGlySerPheLeuGlyAlaArgThrSerAspTrp 104

Qy        348 ATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGGSGTTCCTC 407
              :::||||||||   |||:::||||||||||||   ||||||      |||      |||
Db        105 ValGlyArgArgValThrValValPheAlaAlaAlaIlePheAsnAsnGlySerLeuLeu 124

Qy        408 ATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGC 467
              ||||||||||||||||||||||||||||||||   ||||||||||||   ||||||||||
Db        125 MetGlyPheAlaValAsnTyrAlaMetLeuMetValGlyArgPheValThrGlyIleGly 144

Qy        468 GTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCG 527
```

```
Db        145 ValGlyTyrAlaIleMetValAlaProValTyrThrProGluValSerProAlaSerAla 164

Qy        528 CGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTAC 587
              ||||||||||||||||||||  ||||||||||||||  ||||||||||||||||||
Db        165 ArgGlyPheLeuThrSerPheThrGluValPheIleAsnValGlyIleLeuLeuGlyTyr 184

Qy        588 GTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGC 647
              ||||||||||||||||||:::|||||||||||:::|||  ||||||:::|||||||||
Db        185 ValSerAsnTyrAlaPheAlaArgLeuProLeuHisLeuSerTrpArgValMetLeuGly 204

Qy        648 ATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCG 707
              ||||||||  ||||||  ||||||||||||||||  ||||||||||||||||||||
Db        205 IleGlyAlaValProSerAlaLeuLeuAlaLeuMetValPheGlyMetProGluSerPro 224

Qy        708 CGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCC 767
              |||||||||||||||||||||||||||||||:::  ||||||  ||||||||
Db        225 ArgTrpLeuValMetLysGlyArgLeuAlaAspAlaArgAlaValLeuAlaLysThrSer 244

Qy        768 GACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCT 827
              ||||||  ||||||||  ||||||||      ||||||||||||||||||||||
Db        245 AspThrProGluGluAlaValGluArgLeuAspGlnIleLysAlaAlaAlaGlyIlePro 264

Qy        828 GAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGG 887
                ||||||||||||||||||||  :::||||||        |||  ||||||:::
Db        265 ArgGluLeuAspGlyAspValValValMetProLysThrLysGlyGlyGlnGluLysGln 284

Qy        888 GTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGG 947
              ||||||||||||||||  ||||||||||||||||||||||||||||||:::
Db        285 ValTrpLysGluLeuIlePheSerProThrProAlaMetArgArgIleLeuLeuAlaAla 304

Qy        948 ATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCT 1007
              :::||||||||||||||||  |||  |||  ||||||||  ||||||||
Db        305 LeuGlyIleHisPhePheGlnGlnAlaThrGlySerAspSerValValLeuTyrSerPro 324

Qy       1008 CTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCG 1067
              ||||||:::|||  |||:::|||  |||  |||  ||||||  |||
Db        325 ArgValPheGlnSerAlaGlyIleThrGlyAspAsnHisLeuLeuGlyAlaThrCysAla 344

Qy       1068 TTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGG 1127
                ||||||  |||  ||||||||||||:::|||||||||  :::|||  ||||||
Db        345 MetGlyValMetLysThrLeuPheIleLeuValAlaThrPheGlnLeuAspArgValGly 364

Qy       1128 CGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGGATAATCCTCTCCCTCATCGGCCTCGGC 1187
              ||||||||||||||||  ||||||  |||::::::      ||||||||||||||
Db        365 ArgArgProLeuLeuLeuThrSerThrAlaGlyMetLeuAlaCysLeuIleGlyLeuGly 384

Qy       1188 GCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTA 1247
              ||||||||||||||||:::|||||||||||||:::||||||||||||||||||||
Db        385 ThrGlyLeuThrValValGlyArgHisProAspAlaLysValProTrpAlaIleGlyLeu 404

Qy       1248 AGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACG 1307
                |||  |||  ||||||||||:::|||||||||||||||||||||:::|||
Db        405 CysIleValSerIleLeuAlaTyrValSerPhePheSerIleGlyLeuGlyProLeuThr 424

Qy       1308 TGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGC 1367
              ||||||:::|||||||:::|||||||||:::|||||||||||||||  :::||||||
Db        425 SerValTyrThrSerGluValPheProLeuArgValArgAlaLeuGlyPheAlaLeuGly 444

Qy       1368 GTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAG 1427
                :::  ||||||||||||||||      :::||||||:::|||||||||||||||||
Db        445 ThrSerCysAsnArgValThrSerAlaAlaValSerMetSerPheLeuSerLeuSerLys 464

Qy       1428 GCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTG 1487
              |||||||||||||||||||||||||||||:::||||||||||:::   |||:::
Db        465 AlaIleThrIleGlyGlySerPhePheLeuTyrAlaGlyIleAlaAlaIleGlyTrpIle 484

Qy       1488 TTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTG 1547
              ||||||:::|||:::::|||||||||||||||      |||||||||:::   ||||||
Db        485 PhePhePheThrPheIleProGluThrArgGlyLeuProLeuGluGluIleGlyLysLeu 504

Qy       1548 TTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAGCC---AAGGAGAAGAAGAAG 1604
              ||||||  |||      |||  |||:::  :::  |||  |||:::|||  |||
```

```
Db        505 PheGlyMetThrAspThrAlaValGluAlaGlnAspThrAlaThrLysAspLysAlaLys 524

Qy       1605 GTG 1607
                 |||
Db        525 Val 525

RESULT 6
ABU08339
ID   ABU08339 standard; protein; 529 AA.
XX
AC   ABU08339;
XX
DT   29-MAY-2003  (first entry)
XX
DE   Wheat sugar transport protein #5.
XX
KW   Beta vulgaris-like sugar transport protein; corn; rice; wheat;
KW   plant sugar transport protein; carbohydrate transport; soybean;
KW   carbohydrate distribution; plant.
XX
OS   Triticum aestivum.
XX
PN   US2002178468-A1.
XX
PD   28-NOV-2002.
XX
PF   17-JAN-2002; 2002US-00051902.
XX
PR   24-APR-1998;   98US-0083044P.
PR   14-APR-1999;   99US-00291922.
XX
PA   (ALLE/) ALLEN S M.
PA   (HITZ/) HITZ W D.
PA   (KINN/) KINNEY A J.
PA   (TING/) TINGEY S V.
XX
PI   Allen SM,  Hitz WD,  Kinney AJ,  Tingey SV;
XX
DR   WPI; 2003-340957/32.
DR   N-PSDB; ABX93211.
XX
PT   Novel plant sugar transport proteins and nucleic acid encoding the
PT   protein useful for producing transgenic plants having altered levels of
PT   sugar transport protein.
XX
PS   Claim 10; Fig 2; 56pp; English.
XX
CC   The present invention relates to the isolation of Arabidopsis thaliana-
CC   like or Beta vulgaris-like sugar transport proteins, and the
CC   polynucleotide sequences encoding them. The plant sugar transport
CC   proteins of the invention have been isolated from corn, rice, soybean,
CC   and wheat. The polypeptides of the invention may be used for altering the
CC   level of expression of a sugar transport protein in a host cell, by
CC   transforming a host cell with a chimeric construct encoding all, or a
CC   portion of the sugar transport protein, in sense or antisense
CC   orientation. Particularly, the polypeptides may provide a means to
CC   control carbohydrate transport and distribution in plants. ABU08334-
CC   ABU08339 represent Beta vulgaris-like sugar transport proteins
XX
SQ   Sequence 529 AA;

Alignment Scores:
Pred. No.:                2.5e-133        Length:          529
Score:                    1877.50         Matches:         369
Percent Similarity:       81.38%          Conservative:    55
Best Local Similarity:    70.83%          Mismatches:      92
Query Match:              51.95%          Indels:          5
DB:                       6               Gaps:            3

US-10-051-902A-21 (1-2017) x ABU08339 (1-529)

Qy          60 CCAGAGCGA------GCCTCCTCCTCCTCTGCACCACCGGAGATGGCTTCCGCCGCGCTG 113
               |||||||||      |||                          ||||||||||||||||||
```

```
Db          5 ProGluArgLysGlyAlaGluAspLysGluGluGlySerArgMetAlaSerAlaAlaLeu 24

Qy        114 CCGGAG------GCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGC 167
              ||||||      ||||||   |||:::   |||||||||   ::::::||||||   |||
Db         25 ProGluProGlyAlaValHisProArgAsnLysGlyAsnPheLysTyrAlaPheThrCys 44

Qy        168 GCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGG 227
              |||:::   |||||||||   :::|||:::||||||||||||:::|||||||||||||||
Db         45 AlaLeuCysAlaSerMetAlaThrIleValLeuGlyTyrAspValGlyValMetSerGly 64

Qy        228 GCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATG 287
              ||||||||||||||||||:::|||      |||:::|||   ::::::|||::::::|||
Db         65 AlaSerLeuTyrIleLysArgAspLeuGlnIleThrAspValGlnLeuGluIleMetMet 84

Qy        288 GGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGG 347
              |||||||||::::::|||:::|||||||||||||||      |||||||||||||||||
Db         85 GlyIleLeuSerValTyrAlaLeuIleGlySerPheLeuGlyAlaArgThrSerAspTrp 104

Qy        348 ATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGSGTTCCTC 407
              :::|||||||||   |||:::||||||||||||   ||||||      |||      |||
Db        105 ValGlyArgArgValThrValValPheAlaAlaAlaIlePheAsnAsnGlySerLeuLeu 124

Qy        408 ATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGC 467
              |||||||||||||||||||||||||||||||   ||||||||||||   |||||||||
Db        125 MetGlyPheAlaValAsnTyrAlaMetLeuMetValGlyArgPheValThrGlyIleGly 144

Qy        468 GTGGGCTACGCGCTCATGATCGGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCG 527
              ||||||||||||:::|||:::|||||||||||||||   ||||||||||||||||||||
Db        145 ValGlyTyrAlaIleMetValAlaProValTyrThrProGluValSerProAlaSerAla 164

Qy        528 CGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTAC 587
              ||||||||||||||||||   ||||||||||||||   ||||||||||||||||||||
Db        165 ArgGlyPheLeuThrSerPheThrGluValPheIleAsnValGlyIleLeuLeuGlyTyr 184

Qy        588 GTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGC 647
              |||||||||||||||||:::|||||||||||:::|||   ||||||:::||||||||||
Db        185 ValSerAsnTyrAlaPheAlaArgLeuProLeuHisLeuSerTrpArgValMetLeuGly 204

Qy        648 ATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCG 707
              |||||||||   ||||||   ||||||||||||||||||   ||||||||||||||||||
Db        205 IleGlyAlaValProSerAlaLeuLeuAlaLeuMetValPheGlyMetProGluSerPro 224

Qy        708 CGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCC 767
              |||||||||||||||||||||||||||||||||:::   ||||||   |||||||||
Db        225 ArgTrpLeuValMetLysGlyArgLeuAlaAspAlaArgAlaValLeuAlaLysThrSer 244

Qy        768 GACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCT 827
              ||||||   |||||||||   |||||||||       ||||||||||||||||||||||
Db        245 AspThrProGluGluAlaValGluArgLeuAspGlnIleLysAlaAlaAlaGlyIlePro 264

Qy        828 GAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGG 887
                ||||||||||||||||||||   :::||||||          |||   ||||||:::
Db        265 ArgGluLeuAspGlyAspValValValMetProLysThrLysGlyGlyGlnGluLysGln 284

Qy        888 GTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGG 947
              |||||||||||||||||   |||||||||||||||||||||||||||||||||||:::
Db        285 ValTrpLysGluLeuIlePheSerProThrProAlaMetArgArgIleLeuLeuAlaAla 304

Qy        948 ATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCT 1007
              :::|||||||||||||||||   |||   |||       |||||||||   |||||||||
Db        305 LeuGlyIleHisPhePheGlnGlnAlaThrGlySerAspSerValValLeuTyrSerPro 324

Qy       1008 CTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCG 1067
              ||||||:::|||   |||:::|||   |||   |||   ||||||   |||
Db        325 ArgValPheGlnSerAlaGlyIleThrGlyAspAsnHisLeuLeuGlyAlaThrCysAla 344

Qy       1068 TTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGG 1127
              ||||||   |||   |||||||||||:::|||||||||   :::|||   ||||||
Db        345 MetGlyValMetLysThrLeuPheIleLeuValAlaThrPheGlnLeuAspArgValGly 364

Qy       1128 CGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGC 1187
              |||||||||||||||||   ||||||   |||::::::      |||||||||||||||
Db        365 ArgArgProLeuLeuLeuThrSerThrAlaGlyMetLeuAlaCysLeuIleGlyLeuGly 384

Qy       1188 GCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTA 1247
              |||||||||||||||||||:::|||||||||||||||:::|||||||||||||||||||
Db        385 ThrGlyLeuThrValValGlyArgHisProAspAlaLysValProTrpAlaIleGlyLeu 404

Qy       1248 AGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACG 1307
              |||   |||   |||||||||||:::|||||||||||||||||||||||:::|||
Db        405 CysIleValSerIleLeuAlaTyrValSerPhePheSerIleGlyLeuGlyProLeuThr 424

Qy       1308 TGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGC 1367
              ||||||:::||||||:::|||||||||:::|||||||||||||||   :::||||||
Db        425 SerValTyrThrSerGluValPheProLeuArgValArgAlaLeuGlyPheAlaLeuGly 444

Qy       1368 GTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAG 1427
              :::   ||||||||||||||||      :::||||||:::|||||||||||||||||||
Db        445 ThrSerCysAsnArgValThrSerAlaAlaValSerMetSerPheLeuSerLeuSerLys 464

Qy       1428 GCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTG 1487
              |||||||||||||||||||||||||||||||||||:::|||||||||||||:::   |||:::
Db        465 AlaIleThrIleGlyGlySerPhePheLeuTyrAlaGlyIleAlaAlaIleGlyTrpIle 484

Qy       1488 TTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTG 1547
              ||||||:::|||:::::|||||||||||||||      |||||||||:::   ||||||
Db        485 PhePhePheThrPheIleProGluThrArgGlyLeuProLeuGluGluIleGlyLysLeu 504

Qy       1548 TTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAGCC---AAGGAGAAGAAGAAG 1604
              ||||||   |||      |||   |||:::   :::   |||   |||:::|||   |||
Db        505 PheGlyMetThrAspThrAlaValGluAlaGlnAspThrAlaThrLysAspLysAlaLys 524

Qy       1605 GTG 1607
              |||
Db        525 Val 525
```

RESULT 7
AAU97210
ID AAU97210 standard; protein; 513 AA.
XX
AC AAU97210;
XX
DT 27-AUG-2002 (first entry)
XX
DE Corn Beta vulgaris-like sugar transport protein.
XX
KW Corn; Beta vulgaris-like sugar transport protein; carbohydrate transport;
KW grain filling; annual field crop; plant.
XX
OS Zea mays.
XX
PN US6383776-B1.
XX
PD 07-MAY-2002.
XX
PF 14-APR-1999; 99US-00291922.
XX
PR 24-APR-1998; 98US-0083044P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX
DR WPI; 2002-453364/48.
DR N-PSDB; ABK51971.
XX
PT New nucleic acid encoding plant sugar-transport proteins, useful for
PT preparing transgenic plants with altered carbohydrate distribution.
XX
PS Example 4; Fig 2; 54pp; English.
XX
CC The present invention relates to the isolation of plant polynucleotide
CC sequences encoding an Arabidopsis thaliana-like sugar transport protein
CC or Beta vulgaris-like sugar transport protein. The polynucleotide
CC sequences are useful for altering the level of sugar transport proteins
CC in plants, i.e. for control of carbohydrate transport and distribution in

CC plant cells, e.g. during grain filling of annual field crops (e.g. corn,
CC rice, soybeans, and wheat), and, for studying carbohydrate flows and
CC sugar transport. The polynucleotide sequences can also be used to isolate
CC cDNA sequences and genes that encode homologues of the new proteins. The
CC present sequence represents a corn Beta vulgaris-like sugar transport
CC protein
XX
SQ Sequence 513 AA;

Alignment Scores:

Pred. No.:	5.9e-133	Length:	513
Score:	1872.50	Matches:	372
Percent Similarity:	84.38%	Conservative:	60
Best Local Similarity:	72.66%	Mismatches:	77
Query Match:	51.81%	Indels:	3
DB:	5	Gaps:	2

US-10-051-902A-21 (1-2017) x AAU97210 (1-513)

```
Qy         96 ATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTC 155
Db          1 MetAlaSerAspGluLeuAlaLysAlaValGluProArgLysLysGlyAsnValLysTyr 20

Qy        156 GCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGG 215
Db         21 AlaSerIleCysAlaIleLeuAlaSerMetAlaSerValIleLeuGlyTyrAspIleGly 40

Qy        216 GTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTG 275
Db         41 ValMetSerGlyAlaAlaMetTyrIleLysLysAspLeuAsnIleThrAspValGlnLeu 60

Qy        276 GAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGG 335
Db         61 GluIleLeuIleGlyIleLeuSerLeuTyrSerLeuPheGlySerPheAlaGlyAlaArg 80

Qy        336 ACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCG 395
Db         81 ThrSerAspArgIleGlyArgArgLeuThrValValPheAlaAlaValIlePhePheVal 100

Qy        396 GGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTG 455
Db        101 GlySerLeuLeuMetGlyPheAlaValAsnTyrGlyMetLeuMetAlaGlyArgPheVal 120

Qy        456 GCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCG 515
Db        121 AlaGlyValGlyValGlyTyrGlyGlyMetIleAlaProValTyrThrAlaGluIleSer 140

Qy        516 CCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATC 575
Db        141 ProAlaAlaSerArgGlyPheLeuThrThrPheProGluValPheIleAsnIleGlyIle 160

Qy        576 CTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGC 635
Db        161 LeuLeuGlyTyrLeuSerAsnPheAlaPheAlaArgLeuProLeuHisLeuGlyTrpArg 180

Qy        636 ATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATG 695
Db        181 ValMetLeuAlaIleGlyAlaValProSerGlyLeuLeuAlaLeuLeuValPheCysMet 200

Qy        696 CCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTG 755
Db        201 ProGluSerProArgTrpLeuValLeuLysGlyArgLeuAlaAspAlaArgAlaValLeu 220

Qy        756 GAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCC 815
Db        221 GluLysThrSerAlaThrProGluGluAlaAlaGluArgLeuAlaAspIleLysAlaAla 240

Qy        816 GCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCC---AAGAGAGGGAGC 872
Db        241 AlaGlyIleProLysGlyLeuAspGlyAspValValThrValProGlyLysGluGlnGly 260

Qy        873 GGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGC 932
```

```
Db        261 GlyGlyGluLeuGlnValTrpLysLysLeuIleLeuSerProThrProAlaValArgArg 280

Qy        933 ATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTC 992
Db        281 IleLeuLeuSerAlaValGlyLeuHisPhePheGlnGlnAlaSerGlySerAspSerVal 300

Qy        993 GTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTG 1052
Db        301 ValGlnTyrSerAlaArgLeuPheLysSerAlaGlyIleThrAspAspAsnLysLeuLeu 320

Qy       1053 GGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTC 1112
Db        321 GlyValThrCysAlaValGlyValThrLysThrPhePheIleLeuValAlaThrPheLeu 340

Qy       1113 ATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGGATAATCCTCTCC 1172
Db        341 LeuAspArgAlaGlyArgArgProLeuLeuLeuIleSerThrGlyGlyMetIleValSer 360

Qy       1173 CTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCT 1232
Db        361 LeuIleCysLeuGlySerGlyLeuThrValAlaGlyHisHisProAspThrLysValAla 380

Qy       1233 TGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGC 1292
Db        381 TrpAlaValAlaLeuCysIleAlaSerThrLeuSerTyrIleAlaPhePheSerIleGly 400

Qy       1293 CTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTG 1352
Db        401 LeuGlyProIleThrGlyValTyrThrSerGluIlePheProLeuGlnValArgAlaLeu 420

Qy       1353 GGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTC 1412
Db        421 GlyPheAlaValGlyValAlaSerAsnArgValThrSerAlaValIleSerMetThrPhe 440

Qy       1413 CTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCC 1472
Db        441 LeuSerLeuSerLysAlaIleThrIleGlyGlySerPhePheLeuTyrSerGlyIleAla 460

Qy       1473 GCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAG 1532
Db        461 AlaValAlaTrpValPhePhePheThrCysLeuProGluThrArgGlyArgThrLeuGlu 480

Qy       1533 GAGATGAGCAAGCTGTTCGGC------GACACGGCCGCCGCCTCGGAATCAGACGAGCCA 1586
Db        481 GluMetGlyLysLeuPheGlyMetProAspThrGlyMetAlaGluGluAlaGluAspAla 500

Qy       1587 GCCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCACT 1622
Db        501 AlaAlaLysGluLysValValGluLeuProSerSer 512
```

RESULT 8
ABU08335
ID ABU08335 standard; protein; 513 AA.
XX
AC ABU08335;
XX
DT 29-MAY-2003 (first entry)
XX
DE Corn sugar transport protein #3.
XX
KW Beta vulgaris-like sugar transport protein; corn; rice; wheat;
KW plant sugar transport protein; carbohydrate transport; soybean;
KW carbohydrate distribution; plant.
XX
OS Zea mays.
XX
PN US2002178468-A1.
XX
PD 28-NOV-2002.
XX
PF 17-JAN-2002; 2002US-00051902.
XX
PR 24-APR-1998; 98US-0083044P.

PR 14-APR-1999; 99US-00291922.
XX
PA (ALLE/) ALLEN S M.
PA (HITZ/) HITZ W D.
PA (KINN/) KINNEY A J.
PA (TING/) TINGEY S V.
XX
PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX
DR WPI; 2003-340957/32.
DR N-PSDB; ABX93207.
XX
PT Novel plant sugar transport proteins and nucleic acid encoding the
PT protein useful for producing transgenic plants having altered levels of
PT sugar transport protein.
XX
PS Claim 10; Fig 2; 56pp; English.
XX
CC The present invention relates to the isolation of Arabidopsis thaliana-
CC like or Beta vulgaris-like sugar transport proteins, and the
CC polynucleotide sequences encoding them. The plant sugar transport
CC proteins of the invention have been isolated from corn, rice, soybean,
CC and wheat. The polypeptides of the invention may be used for altering the
CC level of expression of a sugar transport protein in a host cell, by
CC transforming a host cell with a chimeric construct encoding all, or a
CC portion of the sugar transport protein, in sense or antisense
CC orientation. Particularly, the polypeptides may provide a means to
CC control carbohydrate transport and distribution in plants. ABU08334-
CC ABU08339 represent Beta vulgaris-like sugar transport proteins
XX
SQ Sequence 513 AA;

Alignment Scores:

Pred. No.:	5.9e-133	Length:	513
Score:	1872.50	Matches:	372
Percent Similarity:	84.38%	Conservative:	60
Best Local Similarity:	72.66%	Mismatches:	77
Query Match:	51.81%	Indels:	3
DB:	6	Gaps:	2

US-10-051-902A-21 (1-2017) x ABU08335 (1-513)

```
Qy          96 ATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTC 155
               |||||||||      |||   :::||||||   |||:::|||||||||||||||:::::: 
Db           1 MetAlaSerAspGluLeuAlaLysAlaValGluProArgLysLysGlyAsnValLysTyr 20

Qy         156 GCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGG 215
               |||      ||||||||||||||||||||   |||::::::|||||||||||||||||||| 
Db          21 AlaSerIleCysAlaIleLeuAlaSerMetAlaSerValIleLeuGlyTyrAspIleGly 40

Qy         216 GTGATGAGCGGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTG 275
               ||||||||||||||:::::|||||||||||||||   ||||||:::|||   :::::: 
Db          41 ValMetSerGlyAlaAlaMetTyrIleLysLysAspLeuAsnIleThrAspValGlnLeu 60

Qy         276 GAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGGCGG 335
               |||:::|||:::|||||||||:::|||||||||||   ||||||||||      ||| 
Db          61 GluIleLeuIleGlyIleLeuSerLeuTyrSerLeuPheGlySerPheAlaGlyAlaArg 80

Qy         336 ACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCG 395
               |||||||||   ||||||||||||   |||:::|||||||||||||||||||||||||| 
Db          81 ThrSerAspArgIleGlyArgArgLeuThrValValPheAlaAlaValIlePhePheVal 100

Qy         396 GGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTG 455
               |||      |||||||||||||||||||||||   |||||||||   |||||||||||| 
Db         101 GlySerLeuLeuMetGlyPheAlaValAsnTyrGlyMetLeuMetAlaGlyArgPheVal 120

Qy         456 GCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCG 515
               ||||||:::|||||||||||||      |||||||||||||||||||||||||:::||| 
Db         121 AlaGlyValGlyValGlyTyrGlyGlyMetIleAlaProValTyrThrAlaGluIleSer 140

Qy         516 CCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATC 575
               ||||||:::::|||||||||||||||:::|||||||||||||||||||||   |||||| 
Db         141 ProAlaAlaSerArgGlyPheLeuThrThrPheProGluValPheIleAsnIleGlyIle 160

Qy         576 CTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGC 635
               |||||||||||:::||||||:::||||||:::|||||||||||||:::|||||||||||| 
Db         161 LeuLeuGlyTyrLeuSerAsnPheAlaPheAlaArgLeuProLeuHisLeuGlyTrpArg 180

Qy         636 ATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATG 695
               :::||||||   |||||||||   ||||||   |||||||||||||:::|||      ||| 
Db         181 ValMetLeuAlaIleGlyAlaValProSerGlyLeuLeuAlaLeuLeuValPheCysMet 200

Qy         696 CCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTG 755
               |||||||||||||||||||||||:::|||||||||||||||||||:::   |||||| 
Db         201 ProGluSerProArgTrpLeuValLeuLysGlyArgLeuAlaAspAlaArgAlaValLeu 220

Qy         756 GAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCC 815
               ||||||||||||   |||   |||||||||||||||||||||||||||||||||||||| 
Db         221 GluLysThrSerAlaThrProGluGluAlaAlaGluArgLeuAlaAspIleLysAlaAla 240

Qy         816 GCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCC---AAGAGAGGGAGC 872
               ||||||||||||:::   |||||||||||||||||||||||||   |||
Db         241 AlaGlyIleProLysGlyLeuAspGlyAspValValThrValProGlyLysGluGlnGly 260

Qy         873 GGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGC 932
               |||   |||   :::|||||||||:::|||||||||||||||||||||||:::|||||| 
Db         261 GlyGlyGluLeuGlnValTrpLysLysLeuIleLeuSerProThrProAlaValArgArg 280

Qy         933 ATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTC 992
               ||||||||||||   :::|||:::|||||||||||   |||   |||      |||||| 
Db         281 IleLeuLeuSerAlaValGlyLeuHisPhePheGlnGlnAlaSerGlySerAspSerVal 300

Qy         993 GTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTG 1052
               |||   ||||||      :::|||||||||   |||:::|||:::|||      ||| 
Db         301 ValGlnTyrSerAlaArgLeuPheLysSerAlaGlyIleThrAspAspAsnLysLeuLeu 320

Qy        1053 GGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTC 1112
               |||   |||          ||||||||||||      |||||||||:::|||||||||| 
Db         321 GlyValThrCysAlaValGlyValThrLysThrPhePheIleLeuValAlaThrPheLeu 340

Qy        1113 ATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGGATAATCCTCTCC 1172
               :::|||       ||||||||||||||||||||   ||||||||||||:::|||:::||| 
Db         341 LeuAspArgAlaGlyArgArgProLeuLeuLeuIleSerThrGlyGlyMetIleValSer 360

Qy        1173 CTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCT 1232
               ||||||   ||||||:::|||||||||||   |||   |||||||||   |||::: 
Db         361 LeuIleCysLeuGlySerGlyLeuThrValAlaGlyHisHisProAspThrLysValAla 380

Qy        1233 TGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGC 1292
               ||||||:::   |||   |||||||||||||:::|||:::|||||||||||||||||| 
Db         381 TrpAlaValAlaLeuCysIleAlaSerThrLeuSerTyrIleAlaPhePheSerIleGly 400

Qy        1293 CTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTG 1352
               ||||||||||||||   ||||||:::|||||||||||||||||||||||||||||||| 
Db         401 LeuGlyProIleThrGlyValTyrThrSerGluIlePheProLeuGlnValArgAlaLeu 420

Qy        1353 GGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTC 1412
               |||   ||||||||||||:::|||||||||||||||||   |||||||||||||||||| 
Db         421 GlyPheAlaValGlyValAlaSerAsnArgValThrSerAlaValIleSerMetThrPhe 440

Qy        1413 CTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCC 1472
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db         441 LeuSerLeuSerLysAlaIleThrIleGlyGlySerPhePheLeuTyrSerGlyIleAla 460

Qy        1473 GCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAG 1532
               |||:::|||||||||||||||:::|||   |||||||||||||||||||||||||||| 
Db         461 AlaValAlaTrpValPhePhePheThrCysLeuProGluThrArgGlyArgThrLeuGlu 480

Qy        1533 GAGATGAGCAAGCTGTTCGGC------GACACGGCCGCCGCCTCGGAATCAGACGAGCCA 1586
               ||||||   ||||||||||||      ||||||      |||   |||:::::::::: 
Db         481 GluMetGlyLysLeuPheGlyMetProAspThrGlyMetAlaGluGluAlaGluAspAla 500

Qy        1587 GCCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCACT 1622
               |||   ::::::|||   ||||||:::   :::::: 
Db         501 AlaAlaLysGluLysValValGluLeuProSerSer 512
```

RESULT 9
AAU97212
ID AAU97212 standard; protein; 523 AA.
XX
AC AAU97212;
XX
DT 27-AUG-2002 (first entry)
XX
DE Soybean Beta vulgaris-like sugar transport protein.
XX
KW Soybean; Beta vulgaris-like sugar transport protein;
KW carbohydrate transport; grain filling; annual field crop; plant.
XX
OS Glycine max.
XX
PN US6383776-B1.
XX
PD 07-MAY-2002.
XX
PF 14-APR-1999; 99US-00291922.
XX
PR 24-APR-1998; 98US-0083044P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX
DR WPI; 2002-453364/48.
DR N-PSDB; ABK51973.
XX
PT New nucleic acid encoding plant sugar-transport proteins, useful for
PT preparing transgenic plants with altered carbohydrate distribution.
XX
PS Example 4; Fig 2; 54pp; English.
XX
CC The present invention relates to the isolation of plant polynucleotide
CC sequences encoding an Arabidopsis thaliana-like sugar transport protein
CC or Beta vulgaris-like sugar transport protein. The polynucleotide
CC sequences are useful for altering the level of sugar transport proteins
CC in plants, i.e. for control of carbohydrate transport and distribution in
CC plant cells, e.g. during grain filling of annual field crops (e.g. corn,
CC rice, soybeans, and wheat), and, for studying carbohydrate flows and
CC sugar transport. The polynucleotide sequences can also be used to isolate
CC cDNA sequences and genes that encode homologues of the new proteins. The
CC present sequence represents a soybean Beta vulgaris-like sugar transport
CC protein
XX
SQ Sequence 523 AA;

Alignment Scores:

Pred. No.:	7.31e-117	Length:	523
Score:	1660.00	Matches:	331
Percent Similarity:	75.76%	Conservative:	69
Best Local Similarity:	62.69%	Mismatches:	110
Query Match:	45.93%	Indels:	18
DB:	5	Gaps:	4

US-10-051-902A-21 (1-2017) x AAU97212 (1-523)

```
Qy         99 GCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCC 158
              |||         |||   :::         |||||||||:::         ::::::|||
Db         12 AlaHisLysThrLeuGlnAspPheAspProProLysLysArgLysArgAsnLysTyrAla 31

Qy        159 TTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGGGGTG 218
              ||||||||||||:::||||||||||||||||||||||||||||||||||||||||||||||
Db         32 PheAlaCysAlaMetLeuAlaSerMetThrSerIleLeuLeuGlyTyrAspIleGlyVal 51

Qy        219 ATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAG 278
              |||||||||||:::::||||||||:::|||      :::||||||   ::::::|||
Db         52 MetSerGlyAlaAlaIleTyrIleLysArgAspLeuLysValSerAspGluGlnIleGlu 71

Qy        279 GTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACG 338
              :::|||:::||||||:::||||||||||||||||||||||      ||||||||||||
Db         72 IleLeuLeuGlyIleIleAsnLeuTyrSerLeuIleGlySerCysLeuAlaGlyArgThr 91

Qy        339 TCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGG 398
              |||||||||||||||   ||||||||||||||||||||      ||||||||   |||
Db         92 SerAspTrpIleGlyProArgTyrThrIleValPheAlaGlyThrIlePhePheValGly 111

Qy        399 GSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCC 458
                  ||||||||||:::   ||||||:::   ||||||||||||||||||||||||||
Db        112 AlaLeuLeuMetGlyPheSerProAsnTyrSerPheLeuMetPheGlyArgPheValAla 131

Qy        459 GGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCG 518
              |||||||||:::|||||||||||||||||||||||||||||||||||||||||||||||
Db        132 GlyIleGlyIleGlyTyrAlaLeuMetIleAlaProValTyrThrAlaGluValSerPro 151

Qy        519 GCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTG 578
              ||||||:::|||||||||||||||||||||||||||||||||||||   |||||||||
Db        152 AlaSerSerArgGlyPheLeuThrSerPheProGluValPheIleAsnGlyGlyIleLeu 171

Qy        579 CTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATC 638
              :::||||||:::|||||||||||||||||:::|||   |||   :::|||||||||:::
Db        172 IleGlyTyrIleSerAsnTyrAlaPheSerLysLeuThrLeuLysValGlyTrpArgMet 191

Qy        639 ATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCG 698
              |||||||||:::||||||   ||||||||||||||   :::   ||||||   ||||||
Db        192 MetLeuGlyValGlyAlaIleProSerValLeuLeuThrValGlyValLeuAlaMetPro 211

Qy        699 GAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAG 758
              |||||||||||||||||||||:::|||||||||   :::|||:::   ||||||
Db        212 GluSerProArgTrpLeuValMetArgGlyArgLeuGlyGluAlaArgLysValLeuAsn 231

Qy        759 AAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCC 818
              ||||||||||||:::   |||||||||       |||||||||:::||||||   ||||||
Db        232 LysThrSerAspSerLysGluGluAlaGlnLeuArgLeuAlaGluIleLysGlnAlaAla 251

Qy        819 GGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAAC 878
              ||||||||||||      :::   |||||||||   |||   |||:::   :::|||
Db        252 GlyIleProGluSerCysAsnAspAspValValGlnValAsnLysGlnSerAsnGly--- 270

Qy        879 GAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTG 938
                  :::   ||||||||||||||   |||   ||||||||||||:::|||   |||:::
Db        271 ---GluGlyValTrpLysGluLeuPheLeuTyrProThrProAlaIleArgHisIleVal 289

Qy        939 CTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTC 998
              ::::::   :::|||||||||||||||||   |||   |||:::   :::||||||
Db        290 IleAlaAlaLeuGlyIleHisPhePheGlnGlnAlaSerGlyValAspAlaValValLeu 309

Qy        999 TACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGCACC 1058
              ||||||||   :::|||:::      |||:::|||||||||   |||   |||
Db        310 TyrSerProArgIlePheGluLysAlaGlyIleThrAsnAspThrHisLysLeuLeuAla 329

Qy       1059 ACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGAC 1118
              |||         |||      |||   :::|||||||||   |||||||||   :::|||
Db        330 ThrValAlaValGlyPheValLysThrValPheIleLeuAlaAlaThrPheThrLeuAsp 349

Qy       1119 GGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATC 1178
              ||||||||||||||||||||||   |||   ||||||::::::|||||||||:::
Db        350 ArgValGlyArgArgProLeuLeuLeuSerSerValGlyGlyMetValLeuSerLeuLeu 369

Qy       1179 GGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCC 1238
              |||         |||||||||:::      |||   :::   |||:::   ||||||
Db        370 ThrLeuAlaIleSerLeuThrValIle---AspHisSerGluArgLysLeuMetTrpAla 388

Qy       1239 ATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGC 1298
              :::|||   |||||||||      |||||||||||||   |||||||||||   |||
Db        389 ValGlySerSerIleAlaMetValLeuAlaTyrValAlaThrPheSerIleGlyAlaGly 408

Qy       1299 CCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGC 1358
              ||||||||||||||||||||||||||||||||||||:::::||||||   |||
Db        409 ProIleThrTrpValTyrSerSerGluIlePheProLeuArgLeuArgAlaGlnGlyAla 428

Qy       1359 TCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCG 1418
              :::   |||||||||   ||||||   ||||||   |||:::|||||||||||||||||
Db        429 AlaAlaGlyValAlaValAsnArgThrThrSerAlaValValSerMetThrPheLeuSer 448
```

```
Qy       1419 CTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTC 1478
Db        449 LeuThrArgAlaIleThrIleGlyGlyAlaPhePheLeuTyrCysGlyIleAlaThrVal 468

Qy       1479 GCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATG 1538
Db        469 GlyTrpIlePhePheTyrThrValLeuProGluThrArgGlyLysThrLeuGluAspMet 488

Qy       1539 AGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAGCCAAGGAGAAG 1598
Db        489 GluGlySerPheGlyThrPheArgSerLysSerAsnAla------------------Ser 502

Qy       1599 AAGAAGGTGGAAATGGCCGCCACTAACTGATCAAACTAACCGCAAAATCACCAAATCCTA 1658
Db        503 LysAlaValGlu--------------------------AsnGluAsnGlyGlnValAla 513

Qy       1659 AGGGTTTTCTTGCAAAAACGTGTG 1682
Db        514 GlnValGlnLeuGlyThrAsnVal 521
```

RESULT 10
ABU08337
ID ABU08337 standard; protein; 523 AA.
XX
AC ABU08337;
XX
DT 29-MAY-2003 (first entry)
XX
DE Soybean sugar transport protein #3.
XX
KW Beta vulgaris-like sugar transport protein; corn; rice; wheat;
KW plant sugar transport protein; carbohydrate transport; soybean;
KW carbohydrate distribution; plant.
XX
OS Glycine max.
XX
PN US2002178468-A1.
XX
PD 28-NOV-2002.
XX
PF 17-JAN-2002; 2002US-00051902.
XX
PR 24-APR-1998; 98US-0083044P.
PR 14-APR-1999; 99US-00291922.
XX
PA (ALLE/) ALLEN S M.
PA (HITZ/) HITZ W D.
PA (KINN/) KINNEY A J.
PA (TING/) TINGEY S V.
XX
PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX
DR WPI; 2003-340957/32.
DR N-PSDB; ABX93209.
XX
PT Novel plant sugar transport proteins and nucleic acid encoding the
PT protein useful for producing transgenic plants having altered levels of
PT sugar transport protein.
XX
PS Claim 10; Fig 2; 56pp; English.
XX
CC The present invention relates to the isolation of Arabidopsis thaliana-
CC like or Beta vulgaris-like sugar transport proteins, and the
CC polynucleotide sequences encoding them. The plant sugar transport
CC proteins of the invention have been isolated from corn, rice, soybean,
CC and wheat. The polypeptides of the invention may be used for altering the
CC level of expression of a sugar transport protein in a host cell, by
CC transforming a host cell with a chimeric construct encoding all, or a
CC portion of the sugar transport protein, in sense or antisense
CC orientation. Particularly, the polypeptides may provide a means to
CC control carbohydrate transport and distribution in plants. ABU08334-
CC ABU08339 represent Beta vulgaris-like sugar transport proteins
XX
SQ Sequence 523 AA;

Alignment Scores:

Pred. No.:	7.31e-117	Length:	523
Score:	1660.00	Matches:	331
Percent Similarity:	75.76%	Conservative:	69
Best Local Similarity:	62.69%	Mismatches:	110
Query Match:	45.93%	Indels:	18
DB:	6	Gaps:	4

US-10-051-902A-21 (1-2017) x ABU08337 (1-523)

```
Qy         99 GCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCC 158
Db         12 AlaHisLysThrLeuGlnAspPheAspProProLysLysArgLysArgAsnLysTyrAla 31

Qy        159 TTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTG 218
Db         32 PheAlaCysAlaMetLeuAlaSerMetThrSerIleLeuLeuGlyTyrAspIleGlyVal 51

Qy        219 ATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAG 278
Db         52 MetSerGlyAlaAlaIleTyrIleLysArgAspLeuLysValSerAspGluGlnIleGlu 71

Qy        279 GTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGGCGGACG 338
Db         72 IleLeuLeuGlyIleIleAsnLeuTyrSerLeuIleGlySerCysLeuAlaGlyArgThr 91

Qy        339 TCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGG 398
Db         92 SerAspTrpIleGlyProArgTyrThrIleValPheAlaGlyThrIlePhePheValGly 111

Qy        399 GSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCC 458
Db        112 AlaLeuLeuMetGlyPheSerProAsnTyrSerPheLeuMetPheGlyArgPheValAla 131

Qy        459 GGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCG 518
Db        132 GlyIleGlyIleGlyTyrAlaLeuMetIleAlaProValTyrThrAlaGluValSerPro 151

Qy        519 GCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTG 578
Db        152 AlaSerSerArgGlyPheLeuThrSerProGluValPheIleAsnGlyGlyIleLeu 171

Qy        579 CTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATC 638
Db        172 IleGlyTyrIleSerAsnTyrAlaPheSerLysLeuThrLeuLysValGlyTrpArgMet 191

Qy        639 ATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCG 698
Db        192 MetLeuGlyValGlyAlaIleProSerValLeuLeuThrValGlyValLeuAlaMetPro 211

Qy        699 GAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAG 758
Db        212 GluSerProArgTrpLeuValMetArgGlyArgLeuGlyGluAlaArgLysValLeuAsn 231

Qy        759 AAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCC 818
Db        232 LysThrSerAspSerLysGluGluAlaGlnLeuArgLeuAlaGluIleLysGlnAlaAla 251

Qy        819 GGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAAC 878
Db        252 GlyIleProGluSerCysAsnAspAspValValGlnValAsnLysGlnSerAsnGly--- 270

Qy        879 GAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTG 938
Db        271 ---GluGlyValTrpLysGluLeuPheLeuTyrProThrProAlaIleArgHisIleVal 289

Qy        939 CTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTC 998
Db        290 IleAlaAlaLeuGlyIleHisPhePheGlnGlnAlaSerGlyValAspAlaValValLeu 309
```

```
Qy        999 TACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGCACC 1058
Db        310 TyrSerProArgIlePheGluLysAlaGlyIleThrAsnAspThrHisLysLeuLeuAla 329

Qy       1059 ACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGAC 1118
Db        330 ThrValAlaValGlyPheValLysThrValPheIleLeuAlaAlaThrPheThrLeuAsp 349

Qy       1119 GGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATC 1178
Db        350 ArgValGlyArgArgProLeuLeuLeuSerSerValGlyGlyMetValLeuSerLeuLeu 369

Qy       1179 GGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCC 1238
Db        370 ThrLeuAlaIleSerLeuThrValIle---AspHisSerGluArgLysLeuMetTrpAla 388

Qy       1239 ATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGC 1298
Db        389 ValGlySerSerIleAlaMetValLeuAlaTyrValAlaThrPheSerIleGlyAlaGly 408

Qy       1299 CCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGC 1358
Db        409 ProIleThrTrpValTyrSerSerGluIlePheProLeuArgLeuArgAlaGlnGlyAla 428

Qy       1359 TCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCG 1418
Db        429 AlaAlaGlyValAlaValAsnArgThrThrSerAlaValValSerMetThrPheLeuSer 448

Qy       1419 CTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTC 1478
Db        449 LeuThrArgAlaIleThrIleGlyGlyAlaPhePheLeuTyrCysGlyIleAlaThrVal 468

Qy       1479 GCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATG 1538
Db        469 GlyTrpIlePhePheTyrThrValLeuProGluThrArgGlyLysThrLeuGluAspMet 488

Qy       1539 AGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAGCCAAGGAGAAG 1598
Db        489 GluGlySerPheGlyThrPheArgSerLysSerAsnAla-----------------Ser 502

Qy       1599 AAGAAGGTGGAAATGGCCGCCACTAACTGATCAAACTAACCGCAAAATCACCAAATCCTA 1658
Db        503 LysAlaValGlu-------------------------AsnGluAsnGlyGlnValAla 513

Qy       1659 AGGGTTTTCTTGCAAAAACGTGTG 1682
Db        514 GlnValGlnLeuGlyThrAsnVal 521
```

RESULT 11
ABB09681
ID ABB09681 standard; protein; 513 AA.
XX
AC ABB09681;
XX
DT 11-JUN-2002 (first entry)
XX
DE Amino acid sequence of a mannitol transporter designated AgMaT1.
XX
KW Linear polyol; mannitol; mannitol transporter; carbon source; plant;
KW pathogen resistance; salt stress; sorbitol; dulcitol; galactitol;
KW inositol; ribitol; xylitol; AgMaT1.
XX
OS Apium graveolens.
XX
PN WO200204647-A1.
XX
PD 17-JAN-2002.
XX
PF 22-JUN-2001; 2001WO-FR001979.
XX
PR 11-JUL-2000; 2000FR-00009032.
XX
PA (CNRS) CENT NAT RECH SCI.
XX
PI Lemoine RRP, Noiraud NEJ;
XX
DR WPI; 2002-154933/20.
DR N-PSDB; ABL41872.
XX
PT New polyol transporter protein from plants, for selecting transformed
PT cells and for imparting pathogen and salt-stress resistance to plants.
XX
PS Claim 3; Page 39-41; 66pp; French.
XX
CC The present sequence represents AgMaT1, which is a transporter of the
CC linear polyol mannitol. DNA encoding the linear polyol is used to produce
CC a selection system for transformed cells, based on the use of polyols as
CC the only carbon source. It is also used to produce transgenic plants with
CC increased resistance to pathogens and salt stress. The use of a linear
CC polyol for selection eliminates the need for toxic selection reagents
CC such as antibiotics. Also, the linear polyol is not essential for the
CC plant, once selection has been made. The linear polyol has a main chain
CC of 5-8, preferably 6, carbon atoms and can also be sorbitol, dulcitol,
CC galactitol, (myo-)inositol, ribitol, or xylitol
XX
SQ Sequence 513 AA;

Alignment Scores:

Pred. No.:	9.8e-106	Length:	513
Score:	1513.00	Matches:	295
Percent Similarity:	76.77%	Conservative:	85
Best Local Similarity:	59.60%	Mismatches:	109
Query Match:	41.86%	Indels:	6
DB:	5	Gaps:	4

US-10-051-902A-21 (1-2017) x ABB09681 (1-513)

```
Qy        129 CCGAAGAAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACC 188
Db         16 ProLysProLysArgAsn---LysTyrAlaPheAlaCysAlaLeuLeuAlaSerMetAsn 34

Qy        189 TCCATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAG 248
Db         35 SerIleLeuLeuGlyTyrAspThrGlyValLeuSerGlyAlaSerIleTyrIleLysGlu 54

Qy        249 GACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCG 308
Db         55 AspLeuHisPheSerAspValGlnIleGluIleIleIleGlyIleIleAsnIleTyrSer 74

Qy        309 CTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATC 368
Db         75 LeuLeuGlySerAlaIleAlaGlyArgThrSerAspTrpIleGlyArgArgTyrThrMet 94

Qy        369 GTGTTCGCCGCCGTCATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTAC 428
Db         95 ValLeuAlaGlyIleIlePhePheLeuGlyAlaIlePheMetGlyLeuAlaThrAsnPhe 114

Qy        429 GCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATC 488
Db        115 AlaPheLeuMetPheGlyArgPheValAlaGlyIleGlyValGlyTyrAlaMetMetIle 134

Qy        489 GCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTC 548
Db        135 AlaProValTyrThrAlaGluValAlaProSerSerSerArgGlyPheLeuThrSerPhe 154

Qy        549 CCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCC 608
Db        155 ProGluValPheIleAsnSerGlyValLeuLeuGlyTyrValSerAsnPheAlaPheAla 174

Qy        609 CGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTG 668
Db        175 LysCysProLeuTrpLeuGlyTrpArgIleMetLeuGlyIleGlyAlaPheProSerVal 194

Qy        669 CTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGA 728
Db        195 AlaLeuAlaIleIleValLeuTyrMetProGluSerProArgTrpLeuValMetGlnGly 214
```

```
Qy       729 CGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCG 788
Db       215 ArgLeuGlyGluAlaArgThrValLeuGluLysThrSerThrSerLysGluGluAlaHis 234

Qy       789 GAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTG 848
Db       235 GlnArgLeuSerAspIleLysGluAlaAlaGlyIleAspLysAspCysAsnAspAspVal 254

Qy       849 GTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTG 908
Db       255 ValGlnValProLysArg------ThrLysAspGluAlaValTrpLysGluLeuIleLeu 272

Qy       909 TCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAG 968
Db       273 HisProThrLysProValArgHisAlaAlaIleThrGlyIleGlyIleHisPhePheGln 292

Qy       969 CATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGA 1028
Db       293 GlnAlaCysGlyIleAspAlaValValLeuTyrSerProArgIlePheGluLysAlaGly 312

Qy      1029 TTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTT 1088
Db       313 IleLysSerAsnSerLysLysLeuLeuAlaThrIleAlaValGlyValCysLysThrVal 332

Qy      1089 TTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGC 1148
Db       333 PheIleLeuIleSerThrPheGlnLeuAspLysIleGlyArgArgProLeuMetLeuThr 352

Qy      1149 AGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGC 1208
Db       353 SerMetGlyGlyMetValIleAlaLeuPheValLeuAlaGlySerLeuThrValIleAsn 372

Qy      1209 CAG---CACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTC 1265
Db       373 LysSerHisHisThrGlyHis------TrpAlaGlyGlyLeuAlaIlePheThrValTyr 390

Qy      1266 GCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAG 1325
Db       391 AlaPheValSerIlePheSerSerGlyMetGlyProIleAlaTrpValTyrSerSerGlu 410

Qy      1326 ATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTC 1385
Db       411 ValPheProLeuArgLeuArgAlaGlnGlyCysSerIleGlyValAlaValAsnArgGly 430

Qy      1386 ACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGC 1445
Db       431 MetSerGlyIleIleGlyMetThrPheIleSerMetTyrLysAlaMetThrIleGlyGly 450

Qy      1446 AGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTC 1505
Db       451 AlaPheLeuLeuPheAlaValValAlaSerIleGlyTrpValPheMetTyrThrMetPhe 470

Qy      1506 CCGGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGGCCGCC 1565
Db       471 ProGluThrGlnGlyArgAsnLeuGluGluIleGluLeuLeuPheGlySerTyrPheGly 490

Qy      1566 GCCTCGGAATCAGACGAGCCAGCCAAGGAGAAGAAGAAGGTGGAA 1610
Db       491 TrpArgLysThrLeuLysAspLeuLysAlaLysGluAlaAlaGlu 505

RESULT 12
AAG32072
ID   AAG32072 standard; protein; 491 AA.
XX
AC   AAG32072;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 38623.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999;   99US-0121825P.
PR   05-MAR-1999;   99US-0123180P.
PR   09-MAR-1999;   99US-0123548P.
PR   23-MAR-1999;   99US-0125788P.
PR   25-MAR-1999;   99US-0126264P.
PR   29-MAR-1999;   99US-0126785P.
PR   01-APR-1999;   99US-0127462P.
PR   06-APR-1999;   99US-0128234P.
PR   08-APR-1999;   99US-0128714P.
PR   16-APR-1999;   99US-0129845P.
PR   19-APR-1999;   99US-0130077P.
PR   21-APR-1999;   99US-0130449P.
PR   23-APR-1999;   99US-0130510P.
PR   23-APR-1999;   99US-0130891P.
PR   28-APR-1999;   99US-0131449P.
PR   30-APR-1999;   99US-0132048P.
PR   30-APR-1999;   99US-0132407P.
PR   04-MAY-1999;   99US-0132484P.
PR   05-MAY-1999;   99US-0132485P.
PR   06-MAY-1999;   99US-0132486P.
PR   06-MAY-1999;   99US-0132487P.
PR   07-MAY-1999;   99US-0132863P.
PR   11-MAY-1999;   99US-0134256P.
PR   14-MAY-1999;   99US-0134218P.
PR   14-MAY-1999;   99US-0134219P.
PR   14-MAY-1999;   99US-0134221P.
PR   14-MAY-1999;   99US-0134370P.
PR   18-MAY-1999;   99US-0134768P.
PR   19-MAY-1999;   99US-0134941P.
PR   20-MAY-1999;   99US-0135124P.
PR   21-MAY-1999;   99US-0135353P.
PR   24-MAY-1999;   99US-0135629P.
PR   25-MAY-1999;   99US-0136021P.
PR   27-MAY-1999;   99US-0136392P.
PR   28-MAY-1999;   99US-0136782P.
PR   01-JUN-1999;   99US-0137222P.
PR   03-JUN-1999;   99US-0137528P.
PR   04-JUN-1999;   99US-0137502P.
PR   07-JUN-1999;   99US-0137724P.
PR   08-JUN-1999;   99US-0138094P.
PR   10-JUN-1999;   99US-0138540P.
PR   10-JUN-1999;   99US-0138847P.
PR   14-JUN-1999;   99US-0139119P.
PR   16-JUN-1999;   99US-0139452P.
PR   16-JUN-1999;   99US-0139453P.
PR   17-JUN-1999;   99US-0139492P.
PR   18-JUN-1999;   99US-0139454P.
PR   18-JUN-1999;   99US-0139455P.
PR   18-JUN-1999;   99US-0139456P.
PR   18-JUN-1999;   99US-0139457P.
PR   18-JUN-1999;   99US-0139458P.
PR   18-JUN-1999;   99US-0139459P.
PR   18-JUN-1999;   99US-0139460P.
PR   18-JUN-1999;   99US-0139461P.
PR   18-JUN-1999;   99US-0139462P.
PR   18-JUN-1999;   99US-0139463P.
PR   18-JUN-1999;   99US-0139750P.
PR   18-JUN-1999;   99US-0139763P.
PR   21-JUN-1999;   99US-0139817P.
PR   22-JUN-1999;   99US-0139899P.
PR   23-JUN-1999;   99US-0140353P.
PR   23-JUN-1999;   99US-0140354P.
PR   24-JUN-1999;   99US-0140695P.
```

PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:

Pred. No.:	4.35e-96	Length:	491
Score:	1385.50	Matches:	271
Percent Similarity:	72.41%	Conservative:	78
Best Local Similarity:	56.22%	Mismatches:	120
Query Match:	38.34%	Indels:	13
DB:	3	Gaps:	3

US-10-051-902A-21 (1-2017) x AAG32072 (1-491)

```
Qy        150 CGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGAT 209
              :::|||||||||   ||||||||||:::||||||:::   ||||||:::   |||||||||
Db          3 LysPheAlaPheGlyCysAlaIleValAlaSerIleIleSerIleIlePheGlyTyrAsp 22

Qy        210 ATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGG 269
                |||||||||||||||||   ::::::|||:::   |||       |||:::|||
Db         23 ThrGlyValMetSerGlyAlaGlnIlePheIleArgAspAspLeuLysIleAsnAspThr 42

Qy        270 AAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCG 329
              ::::::|||||||||   ||||||||||||||   :::|||:::||||||       |||
Db         43 GlnIleGluValLeuAlaGlyIleLeuAsnLeuCysAlaLeuValGlySerLeuThrAla 62

Qy        330 GGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTC 389
              |||:::|||||||||   ||||||||||||||||||||      :::||||||||||||
Db         63 GlyLysThrSerAspValIleGlyArgArgTyrThrIleAlaLeuSerAlaValIlePhe 82

Qy        390 TTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGC 449
                  |||      |||||||||:::      ||||||   :::||||||   ||||||
```

```
Db        83 LeuValGlySerValLeuMetGlyTyrGlyProAsnTyrProValLeuMetValGlyArg 102

Qy       450 TTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAG 509

Db       103 CysIleAlaGlyValGlyValGlyPheAlaLeuMetIleAlaProValTyrSerAlaGlu 122

Qy       510 GTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTC 569

Db       123 IleSerSerAlaSerHisArgGlyPheLeuThrSerLeuProGluLeuCysIleSerLeu 142

Qy       570 GGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGG 629

Db       143 GlyIleLeuLeuGlyTyrValSerAsnTyrCysPheGlyLysLeuThrLeuLysLeuGly 162

Qy       630 TGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTC 689

Db       163 TrpArgLeuMetLeuGlyIleAlaAlaProSerLeuIleLeuAlaPheGlyIleThr 182

Qy       690 GGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTG 749

Db       183 ArgMetProGluSerProArgTrpLeuValMetGlnGlyArgLeuGluGluAlaLysLys 202

Qy       750 GTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAG 809

Db       203 IleMetValLeuValSerAsnThrGluGluGluAlaGluGluArgPheArgAspIleLeu 222

Qy       810 GCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGG 869

Db       223 ThrAlaAlaGluVal----------------AspValThrGluIleLysGluValGly 236

Qy       870 AGCGGAAACGAGAAGCGG--------------GTGTGGAAGGAGCTCATCCTGTCCCCG 914

Db       237 GlyGlyValLysLysLysAsnHisGlyLysSerValTrpArgGluLeuValIleLysPro 256

Qy       915 ACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCG 974

Db       257 ArgProAlaValArgLeuIleLeuIleAlaAlaValGlyIleHisPhePheGluHisAla 276

Qy       975 TTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACG 1034

Db       277 ThrGlyIleGluAlaValValLeuTyrSerProArgIlePheLysLysAlaGlyValVal 296

Qy      1035 AACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATC 1094

Db       297 SerLysAspLysLeuLeuLeuAlaThrValGlyValGlyLeuThrLysAlaPhePheIle 316

Qy      1095 TTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACG 1154

Db       317 IleIleAlaThrPheLeuLeuAspLysValGlyArgArgLysLeuLeuLeuThrSerThr 336

Qy      1155 GGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCAC 1214

Db       337 GlyGlyMetValPheAlaLeuThrSerLeuAlaValSerLeuThrMetValGlnArgPhe 356

Qy      1215 CCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTC 1274

Db       357 ------GlyArgLeuAlaTrpAlaLeuSerLeuSerIleValSerThrTyrAlaPheVal 374

Qy      1275 GCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCG 1334

Db       375 AlaPhePheSerIleGlyLeuGlyProIleThrTrpValTyrSerSerGluIlePhePro 394

Qy      1335 CTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGC 1394

Db       395 LeuArgLeuArgAlaGlnGlyAlaSerIleGlyValAlaValAsnArgIleMetAsnAla 414

Qy      1395 GTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTC 1454

Db       415 ThrValSerMetSerPheLeuSerMetThrLysAlaIleThrThrGlyGlyValPhePhe 434

Qy      1455 CTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACC 1514

Db       435 ValPheAlaGlyIleAlaValAlaAlaTrpTrpPhePhePheMetLeuProGluThr 454

Qy      1515 CGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAA 1574

Db       455 LysGlyLeuProLeuGluGluMetGluLysLeuPheGlyGlyGlyGlyProArgGlyAsp 474

Qy      1575 TCAGAC 1580

Db       475 ArgAsp 476
```

RESULT 13
AAG32071
ID AAG32071 standard; protein; 508 AA.
XX
AC AAG32071;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 38622.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.

PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:

Pred. No.:	4.4e-96	Length:	508
Score:	1385.50	Matches:	271
Percent Similarity:	72.41%	Conservative:	78
Best Local Similarity:	56.22%	Mismatches:	120
Query Match:	38.34%	Indels:	13
DB:	3	Gaps:	3

```
US-10-051-902A-21 (1-2017) x AAG32071 (1-508)

Qy        150 CGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGAT 209
Db         20 LysPheAlaPheGlyCysAlaIleValAlaSerIleIleSerIleIlePheGlyTyrAsp 39

Qy        210 ATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGG 269
Db         40 ThrGlyValMetSerGlyAlaGlnIlePheIleArgAspAspLeuLysIleAsnAspThr 59

Qy        270 AAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCG 329
Db         60 GlnIleGluValLeuAlaGlyIleLeuAsnLeuCysAlaLeuValGlySerLeuThrAla 79

Qy        330 GGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTC 389
Db         80 GlyLysThrSerAspValIleGlyArgArgTyrThrIleAlaLeuSerAlaValIlePhe 99

Qy        390 TTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGC 449
Db        100 LeuValGlySerValLeuMetGlyTyrGlyProAsnTyrProValLeuMetValGlyArg 119

Qy        450 TTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAG 509
Db        120 CysIleAlaGlyValGlyValGlyPheAlaLeuMetIleAlaProValTyrSerAlaGlu 139

Qy        510 GTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTC 569
Db        140 IleSerSerAlaSerHisArgGlyPheLeuThrSerLeuProGluLeuCysIleSerLeu 159

Qy        570 GGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGG 629
Db        160 GlyIleLeuLeuGlyTyrValSerAsnTyrCysPheGlyLysLeuThrLeuLysLeuGly 179

Qy        630 TGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTC 689
Db        180 TrpArgLeuMetLeuGlyIleAlaAlaProSerLeuIleLeuAlaPheGlyIleThr 199

Qy        690 GGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTG 749
Db        200 ArgMetProGluSerProArgTrpLeuValMetGlnGlyArgLeuGluGluAlaLysLys 219

Qy        750 GTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAG 809
Db        220 IleMetValLeuValSerAsnThrGluGluGluAlaGluGluArgPheArgAspIleLeu 239

Qy        810 GCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGG 869
Db        240 ThrAlaAlaGluVal----------------AspValThrGluIleLysGluValGly 253

Qy        870 AGCGGAAACGAGAAGCGG--------------GTGTGGAAGGAGCTCATCCTGTCCCCG 914
Db        254 GlyGlyValLysLysLysAsnHisGlyLysSerValTrpArgGluLeuValIleLysPro 273

Qy        915 ACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCG 974
Db        274 ArgProAlaValArgLeuIleLeuIleAlaAlaValGlyIleHisPhePheGluHisAla 293

Qy        975 TTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACG 1034
Db        294 ThrGlyIleGluAlaValValLeuTyrSerProArgIlePheLysLysAlaGlyValVal 313

Qy       1035 AACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATC 1094
Db        314 SerLysAspLysLeuLeuLeuAlaThrValGlyValGlyLeuThrLysAlaPhePheIle 333

Qy       1095 TTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACG 1154
Db        334 IleIleAlaThrPheLeuLeuAspLysValGlyArgArgLysLeuLeuLeuThrSerThr 353

Qy       1155 GGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCAC 1214

Db        354 GlyGlyMetValPheAlaLeuThrSerLeuAlaValSerLeuThrMetValGlnArgPhe 373

Qy       1215 CCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTC 1274
Db        374 ------GlyArgLeuAlaTrpAlaLeuSerLeuSerIleValSerThrTyrAlaPheVal 391

Qy       1275 GCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCG 1334
Db        392 AlaPhePheSerIleGlyLeuGlyProIleThrTrpValTyrSerSerGluIlePhePro 411

Qy       1335 CTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGC 1394
Db        412 LeuArgLeuArgAlaGlnGlyAlaSerIleGlyValAlaValAsnArgIleMetAsnAla 431

Qy       1395 GTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTC 1454
Db        432 ThrValSerMetSerPheLeuSerMetThrLysAlaIleThrThrGlyGlyValPhePhe 451

Qy       1455 CTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACC 1514
Db        452 ValPheAlaGlyIleAlaValAlaAlaTrpTrpPhePhePhePheMetLeuProGluThr 471

Qy       1515 CGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAA 1574
Db        472 LysGlyLeuProLeuGluGluMetGluLysLeuPheGlyGlyGlyGlyProArgGlyAsp 491

Qy       1575 TCAGAC 1580
Db        492 ArgAsp 493

RESULT 14
AAG32073
ID   AAG32073 standard; protein; 466 AA.
XX
AC   AAG32073;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 38624.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999;   99US-0121825P.
PR   05-MAR-1999;   99US-0123180P.
PR   09-MAR-1999;   99US-0123548P.
PR   23-MAR-1999;   99US-0125788P.
PR   25-MAR-1999;   99US-0126264P.
PR   29-MAR-1999;   99US-0126785P.
PR   01-APR-1999;   99US-0127462P.
PR   06-APR-1999;   99US-0128234P.
PR   08-APR-1999;   99US-0128714P.
PR   16-APR-1999;   99US-0129845P.
PR   19-APR-1999;   99US-0130077P.
PR   21-APR-1999;   99US-0130449P.
PR   23-APR-1999;   99US-0130510P.
PR   23-APR-1999;   99US-0130891P.
PR   28-APR-1999;   99US-0131449P.
PR   30-APR-1999;   99US-0132048P.
PR   30-APR-1999;   99US-0132407P.
PR   04-MAY-1999;   99US-0132484P.
PR   05-MAY-1999;   99US-0132485P.
PR   06-MAY-1999;   99US-0132486P.
PR   06-MAY-1999;   99US-0132487P.
```

PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.

PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:

Pred. No.:	6.91e-90	Length:	466
Score:	1303.50	Matches:	256
Percent Similarity:	71.90%	Conservative:	74
Best Local Similarity:	55.77%	Mismatches:	116
Query Match:	36.07%	Indels:	13
DB:	3	Gaps:	3

US-10-051-902A-21 (1-2017) x AAG32073 (1-466)

```
Qy        219 ATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAG 278
              ||||||||||||   ::::::|||:::   |||      |||:::|||   ::::::|||
Db          1 MetSerGlyAlaGlnIlePheIleArgAspAspLeuLysIleAsnAspThrGlnIleGlu 20

Qy        279 GTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACG 338
              ||||||   ||||||||||||||||   :::|||:::||||||      ||||||:::|||
Db         21 ValLeuAlaGlyIleLeuAsnLeuCysAlaLeuValGlySerLeuThrAlaGlyLysThr 40

Qy        339 TCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGG 398
              ||||||   ||||||||||||||||||||      :::||||||||||||      |||
Db         41 SerAspValIleGlyArgArgTyrThrIleAlaLeuSerAlaValIlePheLeuValGly 60

Qy        399 GSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCC 458
              ||||||||:::      ||||||   :::||||||   ||||||   :::|||
Db         61 SerValLeuMetGlyTyrGlyProAsnTyrProValLeuMetValGlyArgCysIleAla 80

Qy        459 GGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCG 518
              |||:::|||||||||:::|||||||||||||||||||||||:::||||||:::|||
Db         81 GlyValGlyValGlyPheAlaLeuMetIleAlaProValTyrSerAlaGluIleSerSer 100

Qy        519 GCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTG 578
              ||||||   ||||||||||||||||||   ||||||:::   |||:::   |||||||||
Db        101 AlaSerHisArgGlyPheLeuThrSerLeuProGluLeuCysIleSerLeuGlyIleLeu 120

Qy        579 CTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATC 638
              ||||||||||||||||||||   |||   :::|||   |||   ||||||||||:::
Db        121 LeuGlyTyrValSerAsnTyrCysPheGlyLysLeuThrLeuLysLeuGlyTrpArgLeu 140

Qy        639 ATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCG 698
              ||||||||||   |||   ||||||::::::||||||      :::      ||||||
Db        141 MetLeuGlyIleAlaAlaProSerLeuIleLeuAlaPheGlyIleThrArgMetPro 160

Qy        699 GAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAG 758
              ||||||||||||||||||||||:::|||||||||   :::||||||   ::::::
Db        161 GluSerProArgTrpLeuValMetGlnGlyArgLeuGluGluAlaLysLysIleMetVal 180

Qy        759 AAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCC 818
              |||:::|||   |||||||||   ||||||      ||||||      ||||||
Db        181 LeuValSerAsnThrGluGluGluAlaGluGluArgPheArgAspIleLeuThrAlaAla 200

Qy        819 GGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAAC 878
              :::                    ||||||      :::   :::   |||   |||
Db        201 GluVal------------------AspValThrGluIleLysGluValGlyGlyGlyVal 214

Qy        879 GAGAAGCGG--------------GTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCC 923
              :::|||:::              |||||||:::||||||:::::   |||   ||||||
Db        215 LysLysLysAsnHisGlyLysSerValTrpArgGluLeuValIleLysProArgProAla 234

Qy        924 ATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATT 983
```

```
              :::|||   ||||||::::::   :::||||||||||||||||:::||||||   ||||||
Db        235 ValArgLeuIleLeuIleAlaAlaValGlyIleHisPhePheGluHisAlaThrGlyIle 254

Qy        984 CACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAA 1043
                 :::||||||   |||||||||   :::||||||      |||:::   :::
Db        255 GluAlaValValLeuTyrSerProArgIlePheLysLysAlaGlyValValSerLysAsp 274

Qy       1044 CACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCG 1103
                   |||      |||          |||:::||||||     ||||||:::::|||
Db        275 LysLeuLeuLeuAlaThrValGlyValGlyLeuThrLysAlaPhePheIleIleIleAla 294

Qy       1104 ACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATA 1163
              ||||||   :::|||   ||||||||||   |||||||||   ||||||||||:::
Db        295 ThrPheLeuLeuAspLysValGlyArgArgLysLeuLeuLeuThrSerThrGlyGlyMet 314

Qy       1164 ATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCC 1223
              :::   :::|||      |||          ||||||:::|||   :::
Db        315 ValPheAlaLeuThrSerLeuAlaValSerLeuThrMetValGlnArgPhe------Gly 332

Qy       1224 AAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTC 1283
              ::::::   ||||||:::   |||||||||   ||||||   |||:::||||||||||||
Db        333 ArgLeuAlaTrpAlaLeuSerLeuSerIleValSerThrTyrAlaPheValAlaPhePhe 352

Qy       1284 TCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTG 1343
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||:::::
Db        353 SerIleGlyLeuGlyProIleThrTrpValTyrSerSerGluIlePheProLeuArgLeu 372

Qy       1344 CGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCC 1403
              ||||||   |||   |||:::||||||||||   ||||||:::   :::      :::|||
Db        373 ArgAlaGlnGlyAlaSerIleGlyValAlaValAsnArgIleMetAsnAlaThrValSer 392

Qy       1404 ATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCC 1463
              |||:::|||||||||:::::|||||||||||   ||||||   ||||||:::::::::
Db        393 MetSerPheLeuSerMetThrLysAlaIleThrThrGlyGlyValPhePheValPheAla 412

Qy       1464 GGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGG 1523
              |||||||||      ||||||   ||||||:::      ||||||||||||:::|||
Db        413 GlyIleAlaValAlaAlaTrpTrpPhePhePhePheMetLeuProGluThrLysGlyLeu 432

Qy       1524 ACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGAC 1580
              ||||||||||||   ||||||||||||                        :::   |||
Db        433 ProLeuGluGluMetGluLysLeuPheGlyGlyGlyGlyProArgGlyAspArgAsp 451
```

RESULT 15
ADA48320
ID ADA48320 standard; protein; 333 AA.
XX
AC ADA48320;
XX
DT 20-NOV-2003 (first entry)
XX
DE Rice protein conferring disease resistance in plants.
XX
KW disease resistance; pathogen tolerance; plant pathogen; plant; rice.
XX
OS Oryza sativa.
XX
PN WO2003000906-A2.
XX
PD 03-JAN-2003.
XX
PF 21-JUN-2002; 2002WO-IB002453.
XX
PR 22-JUN-2001; 2001US-0300112P.
PR 26-SEP-2001; 2001US-0352277P.
PR 22-MAR-2002; 2002US-0366535P.
XX
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Glazebrook J, Briggs S, Cooper B, Goff SA, Moughamer T;
PI Katagiri F, Kreps J, Provart N, Ricke D, Zhu T;
XX
DR WPI; 2003-184052/18.

DR N-PSDB; ADA48319.
XX
PT New polynucleotide comprising a plant nucleotide sequence having an open
PT reading frame that encodes a polypeptide associated with disease
PT resistance, useful for conferring resistance or tolerance to a plant
PT pathogen.
XX
PS Claim 10; SEQ ID NO 390; 299pp; English.
XX
CC The invention relates to a novel isolated polynucleotide comprising a
CC plant nucleotide sequence having an open reading frame that encodes a
CC polypeptide associated with disease resistance or its fragment having
CC substantially the same activity as the full-length polypeptide. The
CC polynucleotide of the invention is useful for conferring resistance or
CC tolerance to a plant pathogen. The present sequence represents a protein
CC conferring disease resistance used in the invention.
XX
SQ Sequence 333 AA;

Alignment Scores:

Pred. No.:	6.74e-54	Length:	333
Score:	827.50	Matches:	185
Percent Similarity:	48.28%	Conservative:	53
Best Local Similarity:	37.53%	Mismatches:	82
Query Match:	22.90%	Indels:	173
DB:	6	Gaps:	5

US-10-051-902A-21 (1-2017) x ADA48320 (1-333)

```
Qy         72 TCCTCCTCCTCTGCACCACCGGAGATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCG 131
              ::::::||||||   |||         |||:::|||:::||||||
Db         11 AlaAlaSerSerLysProAlaThrAlaAlaAlaAlaSerLeuProArg------------ 26

Qy        132 AAGAAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCC 191
                           ::::::   |||   ||||||:::|||||||||||||||
Db         27 ------------ArgAsnLysTyrProPhePheCysAlaValLeuAlaSerMetThr*** 42

Qy        192 ATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGAC 251
                                 |||       |||   |||                  :::|||
Db         43 ---------------ArgAlaGlyArgArgSerSerAla--------------GluAsp 52

Qy        252 TTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTC 311
                 :::||||||   ::::::|||||||||   |||   :::|||:::|||||||||
Db         53 LeuGlyValSerAspAlaGlnIleGluValLeuSerGlyAlaIleAsnIleTyrSerLeu 72

Qy        312 ATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTG 371
              :::|||:::       ||||||   |||||||||   :::|||||||||   |||||||||
Db         73 ValGlyAlaLeuLeuAlaGlyTrpThrSerAspArgLeuGlyArgArgLeuThrIleVal 92

Qy        372 TTCGCCGCCGTCATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCC 431
                        |||   ||||||      :::|||      |||      ||||||
Db         93 LeuThrAsnGlyPhePheLeuAlaGlyProLeuValMetSerLeuAlaGlyGlyTyrAla 112

Qy        432 ATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCG 491
                 ||||||   ||||||||||||||||||||||||||||||||||:::||||||
Db        113 AlaLeuMetAlaGlyArgPheValAlaGlyIleGlyValGlyTyrAlaLeuValIleAla 132

Qy        492 CCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCG 551
              |||||||||   ||||||:::||||||||||||:::|||||||   |||:::|||   |||
Db        133 ProValTyrAlaAlaGluIleSerProAlaSerSerArgGlyLeuLeuSerSerLeuPro 152

Qy        552 GAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGC 611
              |||   |||||||||   |||::::::|||   ||||||||||||:::|||||||||
Db        153 Glu---PheIleAsnGlyGlyValMetLeuSerTyrValSerAsnPheAlaPheSerGly 171

Qy        612 TTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTG 671
              ||||||::::::|||   ||||||:::|||         |||      |||:::|||
Db        172 LeuProValHisLeuSerTrpArgLeuMetPheAlaAlaGlyValValProThrValPhe 191

Qy        672 CTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGC 731
              ||||||      ||||||   ||||||||||||||||||||||   ||||||||||||
Db        192 LeuAlaAlaGlyValLeuThrMetProGluSerProArgTrpLeuAlaMetLysGlyArg 211
```

```
Qy        732 CTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAG 791
                    :::|||:::||||||||||::::::||||||||||||      ||||||   :::
Db        212 ArgGlyGluAlaArgValValLeuAspArgThrSerAspThrProAlaGluAlaGluGln 231

Qy        792 CGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTG 851
                  ||||||
Db        232 ArgLeu---------------------------------------------------- 233

Qy        852 ACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCC 911
Db        233 ---------------------------------------------------------- 233

Qy        912 CCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCAT 971
Db        233 ---------------------------------------------------------- 233

Qy        972 GCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTA 1031
Db        233 ---------------------------------------------------------- 233

Qy       1032 ACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTC 1091
Db        233 ---------------------------------------------------------- 233

Qy       1092 ATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGC 1151
Db        233 ---------------------------------------------------------- 233

Qy       1152 ACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAG 1211
Db        233 ---------------------------------------------------------- 233

Qy       1212 CACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTAC 1271
                                                                  :::|||:::
Db        234 ----------------------------------------------ValValAlaPhe 237

Qy       1272 GTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTC 1331
                   |||       ||||||:::|||:::||||||:::   ||||||||||||||||||||
Db        238 ValCysAlaPheSerValGlyIleGlyProLeuAlaTrpValTyrSerSerGluIleLeu 257

Qy       1332 CCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGC 1391
                   ||||||:::::|||       |||      :::|||   |||   |||||||||   |||
Db        258 ProLeuArgLeuArgGlyGlnGlyAlaGlyValGlyThrAlaMetAsnArgValValSer 277

Qy       1392 GGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTC 1451
                   ||||||::::::|||||||||:::||||||       |||||||||:::   |||:::|||
Db        278 GlyValValThrMetThrPheIleSerLeuTyrGlyAlaIleThrMetAlaGlyAlaPhe 297

Qy       1452 TTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAG 1511
                   :::||||||:::   |||||||||   ::::::||||||   |||      |||||||||
Db        298 TyrLeuTyrAlaAlaIleAlaAlaAlaSerPheValPheIleTyrAlaCysLeuProGlu 317

Qy       1512 ACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTC 1550
                   ||||||||||||:::||||||:::|||   :::||||||
Db        318 ThrArgGlyArgSerLeuGluAspMetGluGluLeuPhe 330
```

Search completed: June 30, 2004, 18:48:02
Job time : 129 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 30, 2004, 18:42:55 ; Search time 24 Seconds
(without alignments)
8677.465 Million cell updates/sec

Title: US-10-051-902A-21
Perfect score: 3614
Sequence: 1 cttacatgtaagctcgtgcc.....aaaaaaaaaaaaaaaaaaa 2017

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 778828

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10051902/runat_30062004_164724_20438/app_query.fasta_1.2183
-DB=Issued_Patents_AA -QFMT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10051902_@CGN_1_1_27_@runat_30062004_164724_20438 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2570	71.1	510	4	US-09-291-922-22	Sequence 22, Appl
2	2015	55.8	539	4	US-09-291-922-26	Sequence 26, Appl
3	1877.5	52.0	529	4	US-09-291-922-28	Sequence 28, Appl
4	1872.5	51.8	513	4	US-09-291-922-20	Sequence 20, Appl
5	1660	45.9	523	4	US-09-291-922-24	Sequence 24, Appl
6	1589	44.0	549	4	US-09-291-922-30	Sequence 30, Appl
7	682	18.9	167	4	US-09-291-922-18	Sequence 18, Appl
8	672	18.6	488	4	US-10-162-012-46	Sequence 46, Appl
9	591	16.4	514	4	US-09-489-039A-11902	Sequence 11902, A
10	585	16.2	476	4	US-09-489-039A-11933	Sequence 11933, A
11	576	15.9	584	2	US-08-928-692-13	Sequence 13, Appl
12	576	15.9	584	4	US-09-339-972-13	Sequence 13, Appl
13	549	15.2	517	4	US-09-679-686B-18	Sequence 18, Appl
14	536	14.8	514	4	US-09-679-686B-22	Sequence 22, Appl
15	533	14.7	502	4	US-09-679-686B-2	Sequence 2, Appli
16	520	14.4	511	4	US-09-679-686B-12	Sequence 12, Appl
17	514.5	14.2	501	4	US-09-489-039A-11731	Sequence 11731, A
18	507.5	14.0	519	4	US-09-679-686B-24	Sequence 24, Appl
19	505.5	14.0	510	4	US-09-679-686B-19	Sequence 19, Appl
20	500	13.8	729	4	US-09-291-922-29	Sequence 29, Appl
21	499	13.8	518	4	US-09-679-686B-23	Sequence 23, Appl
22	489.5	13.5	504	4	US-09-679-686B-21	Sequence 21, Appl
23	487.5	13.5	518	4	US-09-679-686B-16	Sequence 16, Appl
24	487	13.5	562	4	US-10-162-012-44	Sequence 44, Appl
25	481	13.3	737	4	US-09-291-922-8	Sequence 8, Appli
26	469	13.0	493	2	US-09-031-392-10	Sequence 10, Appl
27	469	13.0	493	3	US-09-299-549-10	Sequence 10, Appl
28	469	13.0	493	4	US-09-610-417-10	Sequence 10, Appl
29	453	12.5	521	4	US-09-489-039A-9549	Sequence 9549, Ap
30	448	12.4	509	2	US-09-031-392-6	Sequence 6, Appli
31	448	12.4	509	3	US-09-299-549-6	Sequence 6, Appli
32	448	12.4	509	4	US-09-610-417-6	Sequence 6, Appli
33	443	12.3	494	2	US-09-031-392-5	Sequence 5, Appli
34	443	12.3	494	3	US-09-299-549-5	Sequence 5, Appli
35	443	12.3	494	4	US-09-610-417-5	Sequence 5, Appli
36	439.5	12.2	492	2	US-08-355-844-3	Sequence 3, Appli
37	439.5	12.2	492	5	PCT-US95-16126-3	Sequence 3, Appli
38	438.5	12.1	488	2	US-08-928-692-11	Sequence 11, Appl
39	438.5	12.1	488	4	US-09-339-972-11	Sequence 11, Appl
40	437	12.1	747	4	US-09-291-922-2	Sequence 2, Appli
41	434.5	12.0	323	4	US-09-134-001C-4635	Sequence 4635, Ap
42	422	11.7	524	2	US-08-928-692-12	Sequence 12, Appl
43	422	11.7	524	4	US-09-339-972-12	Sequence 12, Appl
44	420.5	11.6	534	2	US-09-031-392-4	Sequence 4, Appli
45	420.5	11.6	534	3	US-09-299-549-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-09-291-922-22
; Sequence 22, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 22
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (102)
US-09-291-922-22

Alignment Scores:

Pred. No.:	3.19e-189	Length:	510
Score:	2570.00	Matches:	510
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	71.11%	Indels:	0
DB:	4	Gaps:	0

US-10-051-902A-21 (1-2017) x US-09-291-922-22 (1-510)

Appl.

```
Qy        96 ATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTC 155
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 MetAlaSerAlaAlaLeuProGluAlaValAlaProLysLysLysGlyAsnValArgPhe 20

Qy       156 GCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGG 215
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        21 AlaPheAlaCysAlaIleLeuAlaSerMetThrSerIleLeuLeuGlyTyrAspIleGly 40

Qy       216 GTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTG 275
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        41 ValMetSerGlyAlaSerLeuTyrIleLysLysAspPheAsnIleSerAspGlyLysVal 60

Qy       276 GAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGG 335
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 GluValLeuMetGlyIleLeuAsnLeuTyrSerLeuIleGlySerPheAlaAlaGlyArg 80

Qy       336 ACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCG 395
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        81 ThrSerAspTrpIleGlyArgArgTyrThrIleValPheAlaAlaValIlePhePheAla 100

Qy       396 GGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTG 455
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       101 Gly***PheLeuMetGlyPheAlaValAsnTyrAlaMetLeuMetPheGlyArgPheVal 120

Qy       456 GCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCG 515
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 AlaGlyIleGlyValGlyTyrAlaLeuMetIleAlaProValTyrThrAlaGluValSer 140

Qy       516 CCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATC 575
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       141 ProAlaSerAlaArgGlyPheLeuThrSerPheProGluValPheIleAsnPheGlyIle 160

Qy       576 CTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGC 635
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       161 LeuLeuGlyTyrValSerAsnTyrAlaPheSerArgLeuProLeuAsnLeuGlyTrpArg 180

Qy       636 ATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATG 695
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 IleMetLeuGlyIleGlyAlaAlaProSerValLeuLeuAlaLeuMetValLeuGlyMet 200

Qy       696 CCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTG 755
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       201 ProGluSerProArgTrpLeuValMetLysGlyArgLeuAlaAspAlaLysValValLeu 220

Qy       756 GAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCC 815
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       221 GluLysThrSerAspThrAlaGluGluAlaAlaGluArgLeuAlaAspIleLysAlaAla 240

Qy       816 GCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGA 875
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 AlaGlyIleProGluGluLeuAspGlyAspValValThrValProLysArgGlySerGly 260

Qy       876 AACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATC 935
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       261 AsnGluLysArgValTrpLysGluLeuIleLeuSerProThrProAlaMetArgArgIle 280

Qy       936 CTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTC 995
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       281 LeuLeuSerGlyIleGlyIleHisPhePheGlnHisAlaLeuGlyIleHisSerValVal 300

Qy       996 TTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGC 1055
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 PheTyrSerProLeuValPheLysSerProGlyLeuThrAsnAspLysHisPheLeuGly 320

Qy      1056 ACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATC 1115
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       321 ThrThrTrpProPheGlyValThrLysArgLeuPheIleLeuLeuAlaThrPhePheIle 340

Qy      1116 GACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTC 1175
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       341 AspGlyValGlyArgArgProLeuLeuLeuGlySerThrGlyGlyIleIleLeuSerLeu 360

             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 IleGlyLeuGlyAlaGlyLeuThrValValGlyGlnHisProAspAlaLysIleProTrp 380

Qy      1236 GCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTT 1295
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       381 AlaIleGlyLeuSerIleAlaSerThrLeuAlaTyrValAlaPhePheSerIleGlyLeu 400

Qy      1296 GGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGC 1355
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       401 GlyProIleThrTrpValTyrSerSerGluIlePheProLeuGlnValArgAlaLeuGly 420

Qy      1356 TGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTG 1415
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 CysSerLeuGlyValAlaAlaAsnArgValThrSerGlyValIleSerMetThrPheLeu 440

Qy      1416 TCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCG 1475
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       441 SerLeuSerLysAlaIleThrIleGlyGlySerPhePheLeuTyrSerGlyIleAlaAla 460

Qy      1476 CTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAG 1535
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       461 LeuAlaTrpValPhePheTyrThrTyrLeuProGluThrArgGlyArgThrLeuGluGlu 480

Qy      1536 ATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAGCCAAGGAG 1595
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 MetSerLysLeuPheGlyAspThrAlaAlaAlaSerGluSerAspGluProAlaLysGlu 500

Qy      1596 AAGAAGAAGGTGGAAATGGCCGCCACTAAC 1625
             ||||||||||||||||||||||||||||||
Db       501 LysLysLysValGluMetAlaAlaThrAsn 510
```

RESULT 2
US-09-291-922-26
; Sequence 26, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 26
; LENGTH: 539
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-09-291-922-26

Alignment Scores:

Pred. No.:	1.34e-146	Length:	539
Score:	2015.00	Matches:	408
Percent Similarity:	84.49%	Conservative:	44
Best Local Similarity:	76.26%	Mismatches:	73
Query Match:	55.76%	Indels:	10
DB:	4	Gaps:	4

US-10-051-902A-21 (1-2017) x US-09-291-922-26 (1-539)

```
Qy        48 ACTGTACACGGCCCAGAGCGA-----------GCCTCCTCCTCCTCTGCACCACCGGAG 95
             |||      ||||||   |||           |||:::|||     |||
Db         6 ThrGlnGlyGlyProArgArgHisAsnProGlnAlaAlaSerArgGlyAlaSerSerThr 25

Qy        96 ATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTC 155
             |||      ||||||||||||   ||||||   ||||||||||||||||||||||||||
Db        26 MetAspArgAlaAlaLeuProAlaAlaValGluProLysLysLysGlyAsnValArgPhe 45
```

```
Db         46 AlaPheAlaCysAlaIleLeuAlaSerMetThrSerIleLeuLeuGlyTyrAspIleGly 65

Qy        216 GTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTG 275
Db         66 ValMetSerGlyAlaSerLeuTyrIleGlnLysAspLeuLysIleAsnAspThrGlnLeu 85

Qy        276 GAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGG 335
Db         86 GluValLeuMetGlyIleLeuAsnValTyrSerLeuIleGlySerPheAlaAlaGlyArg 105

Qy        336 ACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCG 395
Db        106 ThrSerAspTrpIleGlyArgArgPheThrIleValPheAlaAlaValIlePhePheAla 125

Qy        396 GGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTG 455
Db        126 GlyAlaLeuIleMetGlyPheSerValAsnTyrAlaMetLeuMetPheGlyArgPheVal 145

Qy        456 GCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCG 515
Db        146 AlaGlyIleGlyValGlyTyrAlaLeuMetIleAlaProValAsnThrGlyGluValSer 165

Qy        516 CCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATC 575
Db        166 ProAlaSerAlaArgGlyValLeuThrSerPheProGluValPheIleAsnPheGlyIle 185

Qy        576 CTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGC 635
Db        186 LeuLeuGlyTyrValSerAsnPheAlaPheAlaArgLeuSerLeuArgLeuGlyTrpArg 205

Qy        636 ATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATG 695
Db        206 IleMetLeuGlyIleGlyAlaValProSerValLeuLeuAlaPheMetValLeuGlyMet 225

Qy        696 CCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTG 755
Db        226 ProGluSerProArgTrpLeuValMetLysGlyArgLeuAlaAspAlaLysValValLeu 245

Qy        756 GAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCC 815
Db        246 AlaLysThrSerAspThrProGluGluAlaAlaGluArgIleAlaAspIleLysThrAla 265

Qy        816 GCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAG---AGAGGGAGC 872
Db        266 AlaGlyIleProLeuGlyLeuAspGlyAspValValProValProLysAsnLysGlySer 285

Qy        873 GGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGC 932
Db        286 SerGluGluLysArgValLeuLysAspLeuIleLeuSerProThrIleAlaMetArgHis 305

Qy        933 ATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTC 992
Db        306 IleLeuIleAlaGlyIleGlyIleHisPhePheGlnGlnSerSerGlyIleAspAlaVal 325

Qy        993 GTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTG 1052
Db        326 ValLeuTyrSerProLeuValPheLysSerAlaGlyIleThrGlyAspSerArgLeuArg 345

Qy       1053 GGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTC 1112
Db        346 GlyThrThrValAlaValGlyAlaThrAsnThrValPheIleLeuValAlaThrPheLeu 365

Qy       1113 ATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCC 1172
Db        366 LeuAspArgIleArgArgArgProLeuValLeuThrSerThrGlyGlyMetLeuValSer 385

Qy       1173 CTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCT 1232
Db        386 LeuValGlyLeuAlaThrGlyLeuThrValIleSerArgHisProAspGluLysIleThr 405

Qy       1233 TGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGC 1292
```

```
Db        406 TrpAlaIleValLeuCysIlePheCysIleMetAlaTyrValAlaPhePheSerIleGly 425

Qy       1293 CTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTG 1352
Db        426 LeuGlyProIleThrTrpValTyrSerSerGluIlePheProLeuHisValArgAlaLeu 445

Qy       1353 GGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTC 1412
Db        446 GlyCysSerLeuGlyValAlaValAsnArgLeuThrSerGlyValIleSerMetThrPhe 465

Qy       1413 CTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCC 1472
Db        466 IleSerLeuSerLysAlaMetThrIleGlyGlyAlaPhePheLeuPheAlaGlyIleAla 485

Qy       1473 GCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAG 1532
Db        486 SerPheAlaTrpValPhePhePheAlaTyrLeuProGluThrArgGlyArgThrLeuGlu 505

Qy       1533 GAGATGAGCAAGCTGTTCGGCGACACGGCC----------GCCGCCTCGGAATCAGAC 1580
Db        506 AspMetSerSerLeuPheGlyAsnThrAlaThrHisLysGlnGlyAlaAlaGluAlaAsp 525

Qy       1581 GAGCCAGCCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCACTAAC 1625
Db        526 AspAspAlaGlyGlu---LysLysValGluMetAlaAlaThrAsn 539
```

RESULT 3
US-09-291-922-28
; Sequence 28, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 28
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-09-291-922-28

Alignment Scores:

Pred. No.:	4.8e-136	Length:	529
Score:	1877.50	Matches:	369
Percent Similarity:	81.38%	Conservative:	55
Best Local Similarity:	70.83%	Mismatches:	92
Query Match:	51.95%	Indels:	5
DB:	4	Gaps:	3

US-10-051-902A-21 (1-2017) x US-09-291-922-28 (1-529)

```
Qy         60 CCAGAGCGA------GCCTCCTCCTCCTCTGCACCACCGGAGATGGCTTCCGCCGCGCTG 113
Db          5 ProGluArgLysGlyAlaGluAspLysGluGluGlySerArgMetAlaSerAlaAlaLeu 24

Qy        114 CCGGAG------GCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGC 167
Db         25 ProGluProGlyAlaValHisProArgAsnLysGlyAsnPheLysTyrAlaPheThrCys 44

Qy        168 GCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGG 227
Db         45 AlaLeuCysAlaSerMetAlaThrIleValLeuGlyTyrAspValGlyValMetSerGly 64

Qy        228 GCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATG 287
```

```
Db         65 AlaSerLeuTyrIleLysArgAspLeuGlnIleThrAspValGlnLeuGluIleMetMet 84

Qy        288 GGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGG 347
Db         85 GlyIleLeuSerValTyrAlaLeuIleGlySerPheLeuGlyAlaArgThrSerAspTrp 104

Qy        348 ATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGSGTTCCTC 407
Db        105 ValGlyArgArgValThrValValPheAlaAlaAlaIlePheAsnAsnGlySerLeuLeu 124

Qy        408 ATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGC 467
Db        125 MetGlyPheAlaValAsnTyrAlaMetLeuMetValGlyArgPheValThrGlyIleGly 144

Qy        468 GTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCG 527
Db        145 ValGlyTyrAlaIleMetValAlaProValTyrThrProGluValSerProAlaSerAla 164

Qy        528 CGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTAC 587
Db        165 ArgGlyPheLeuThrSerPheThrGluValPheIleAsnValGlyIleLeuLeuGlyTyr 184

Qy        588 GTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGC 647
Db        185 ValSerAsnTyrAlaPheAlaArgLeuProLeuHisLeuSerTrpArgValMetLeuGly 204

Qy        648 ATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCG 707
Db        205 IleGlyAlaValProSerAlaLeuLeuAlaLeuMetValPheGlyMetProGluSerPro 224

Qy        708 CGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCC 767
Db        225 ArgTrpLeuValMetLysGlyArgLeuAlaAspAlaArgAlaValLeuAlaLysThrSer 244

Qy        768 GACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCT 827
Db        245 AspThrProGluGluAlaValGluArgLeuAspGlnIleLysAlaAlaAlaGlyIlePro 264

Qy        828 GAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGG 887
Db        265 ArgGluLeuAspGlyAspValValValMetProLysThrLysGlyGlyGlnGluLysGln 284

Qy        888 GTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGG 947
Db        285 ValTrpLysGluLeuIlePheSerProThrProAlaMetArgArgIleLeuLeuAlaAla 304

Qy        948 ATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCT 1007
Db        305 LeuGlyIleHisPhePheGlnGlnAlaThrGlySerAspSerValValLeuTyrSerPro 324

Qy       1008 CTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCG 1067
Db        325 ArgValPheGlnSerAlaGlyIleThrGlyAspAsnHisLeuLeuGlyAlaThrCysAla 344

Qy       1068 TTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGG 1127
Db        345 MetGlyValMetLysThrLeuPheIleLeuValAlaThrPheGlnLeuAspArgValGly 364

Qy       1128 CGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGC 1187
Db        365 ArgArgProLeuLeuLeuThrSerThrAlaGlyMetLeuAlaCysLeuIleGlyLeuGly 384

Qy       1188 GCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTA 1247
Db        385 ThrGlyLeuThrValValGlyArgHisProAspAlaLysValProTrpAlaIleGlyLeu 404

Qy       1248 AGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACG 1307
Db        405 CysIleValSerIleLeuAlaTyrValSerPhePheSerIleGlyLeuGlyProLeuThr 424

Qy       1308 TGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGC 1367
Db        425 SerValTyrThrSerGluValPheProLeuArgValArgAlaLeuGlyPheAlaLeuGly 444

Qy       1368 GTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAG 1427
Db        445 ThrSerCysAsnArgValThrSerAlaAlaValSerMetSerPheLeuSerLeuSerLys 464

Qy       1428 GCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTG 1487
Db        465 AlaIleThrIleGlyGlySerPhePheLeuTyrAlaGlyIleAlaAlaIleGlyTrpIle 484

Qy       1488 TTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTG 1547
Db        485 PhePhePheThrPheIleProGluThrArgGlyLeuProLeuGluGluIleGlyLysLeu 504

Qy       1548 TTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAGCC---AAGGAGAAGAAGAAG 1604
Db        505 PheGlyMetThrAspThrAlaValGluAlaGlnAspThrAlaThrLysAspLysAlaLys 524

Qy       1605 GTG 1607
Db        525 Val 525
```

RESULT 4
US-09-291-922-20
; Sequence 20, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 20
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Zea mays
US-09-291-922-20

Alignment Scores:

Pred. No.:	1.15e-135	Length:	513
Score:	1872.50	Matches:	372
Percent Similarity:	84.38%	Conservative:	60
Best Local Similarity:	72.66%	Mismatches:	77
Query Match:	51.81%	Indels:	3
DB:	4	Gaps:	2

US-10-051-902A-21 (1-2017) x US-09-291-922-20 (1-513)

```
Qy         96 ATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTC 155
Db          1 MetAlaSerAspGluLeuAlaLysAlaValGluProArgLysLysGlyAsnValLysTyr 20

Qy        156 GCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGG 215
Db         21 AlaSerIleCysAlaIleLeuAlaSerMetAlaSerValIleLeuGlyTyrAspIleGly 40

Qy        216 GTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTG 275
Db         41 ValMetSerGlyAlaAlaMetTyrIleLysLysAspLeuAsnIleThrAspValGlnLeu 60

Qy        276 GAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGG 335
Db         61 GluIleLeuIleGlyIleLeuSerLeuTyrSerLeuPheGlySerPheAlaGlyAlaArg 80

Qy        336 ACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCG 395
Db         81 ThrSerAspArgIleGlyArgArgLeuThrValValPheAlaAlaValIlePhePheVal 100
```

```
Qy        396 GGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTG 455
              |||      ||||||||||||||||||||||||   |||||||||   ||||||||||||
Db        101 GlySerLeuLeuMetGlyPheAlaValAsnTyrGlyMetLeuMetAlaGlyArgPheVal 120

Qy        456 GCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCG 515
              ||||||:::|||||||||||||      |||||||||||||||||||||||||||:::|||
Db        121 AlaGlyValGlyValGlyTyrGlyGlyMetIleAlaProValTyrThrAlaGluIleSer 140

Qy        516 CCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATC 575
              ||||||::::::|||||||||||||||:::|||||||||||||||||||||||   ||||||
Db        141 ProAlaAlaSerArgGlyPheLeuThrThrPheProGluValPheIleAsnIleGlyIle 160

Qy        576 CTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGC 635
              |||||||||||:::||||||:::||||||:::|||||||||||||:::|||||||||||||
Db        161 LeuLeuGlyTyrLeuSerAsnPheAlaPheAlaArgLeuProLeuHisLeuGlyTrpArg 180

Qy        636 ATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATG 695
              :::||||||   ||||||||   ||||||   ||||||||||||:::|||      |||
Db        181 ValMetLeuAlaIleGlyAlaValProSerGlyLeuLeuAlaLeuLeuValPheCysMet 200

Qy        696 CCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTG 755
              |||||||||||||||||||||||:::|||||||||||||||||:::   ||||||
Db        201 ProGluSerProArgTrpLeuValLeuLysGlyArgLeuAlaAspAlaArgAlaValLeu 220

Qy        756 GAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCC 815
              ||||||||||||   |||   ||||||||||||||||||||||||||||||||||||||||
Db        221 GluLysThrSerAlaThrProGluGluAlaAlaGluArgLeuAlaAspIleLysAlaAla 240

Qy        816 GCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCC---AAGAGAGGGAGC 872
              ||||||||||||:::   |||||||||||||||||||||||||   |||
Db        241 AlaGlyIleProLysGlyLeuAspGlyAspValValThrValProGlyLysGluGlnGly 260

Qy        873 GGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGC 932
              |||   |||   :::|||||||||:::|||||||||||||||||||||||||:::||||||
Db        261 GlyGlyGluLeuGlnValTrpLysLysLeuIleLeuSerProThrProAlaValArgArg 280

Qy        933 ATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTC 992
              ||||||||||||   :::|||:::||||||||||||   |||   |||      ||||||
Db        281 IleLeuLeuSerAlaValGlyLeuHisPhePheGlnGlnAlaSerGlySerAspSerVal 300

Qy        993 GTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTG 1052
              |||   ||||||      :::|||||||||   |||:::|||:::|||         |||
Db        301 ValGlnTyrSerAlaArgLeuPheLysSerAlaGlyIleThrAspAspAsnLysLeuLeu 320

Qy        1053 GGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTC 1112
               |||   |||          ||||||||||      |||||||||:::||||||||
Db         321 GlyValThrCysAlaValGlyValThrLysThrPhePheIleLeuValAlaThrPheLeu 340

Qy        1113 ATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCC 1172
               :::|||      |||||||||||||||||||||   ||||||||||:::|||:::|||
Db         341 LeuAspArgAlaGlyArgArgProLeuLeuLeuIleSerThrGlyGlyMetIleValSer 360

Qy        1173 CTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCT 1232
               ||||||   ||||||:::||||||||||||   |||   |||||||||   |||:::
Db         361 LeuIleCysLeuGlySerGlyLeuThrValAlaGlyHisHisProAspThrLysValAla 380

Qy        1233 TGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGC 1292
               ||||||:::   |||   |||||||||||||||:::|||:::|||||||||||||||||||
Db         381 TrpAlaValAlaLeuCysIleAlaSerThrLeuSerTyrIleAlaPhePheSerIleGly 400

Qy        1293 CTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTG 1352
               ||||||||||||||   ||||||:::|||||||||||||||||||||||||||||||||||
Db         401 LeuGlyProIleThrGlyValTyrThrSerGluIlePheProLeuGlnValArgAlaLeu 420

Qy        1353 GGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTC 1412
               |||   ::::::||||||||||:::|||||||||||||||   ||||||||||||||||||
Db         421 GlyPheAlaValGlyValAlaSerAsnArgValThrSerAlaValIleSerMetThrPhe 440

Qy        1413 CTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCC 1472
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         441 LeuSerLeuSerLysAlaIleThrIleGlyGlySerPhePheLeuTyrSerGlyIleAla 460

Qy        1473 GCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAG 1532
               |||:::|||||||||||||||:::|||   |||||||||||||||||||||||||||||||
Db         461 AlaValAlaTrpValPhePhePheThrCysLeuProGluThrArgGlyArgThrLeuGlu 480

Qy        1533 GAGATGAGCAAGCTGTTCGGC------GACACGGCCGCCGCCTCGGAATCAGACGAGCCA 1586
               ||||||   ||||||||||||      ||||||      |||   |||:::::::::
Db         481 GluMetGlyLysLeuPheGlyMetProAspThrGlyMetAlaGluGluAlaGluAspAla 500

Qy        1587 GCCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCACT 1622
               |||   ::::::|||   ||||||:::   ::::::
Db         501 AlaAlaLysGluLysValValGluLeuProSerSer 512
```

RESULT 5
US-09-291-922-24
; Sequence 24, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 24
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Glycine max
US-09-291-922-24

Alignment Scores:

Pred. No.:	2.39e-119	Length:	523
Score:	1660.00	Matches:	331
Percent Similarity:	75.76%	Conservative:	69
Best Local Similarity:	62.69%	Mismatches:	110
Query Match:	45.93%	Indels:	18
DB:	4	Gaps:	4

US-10-051-902A-21 (1-2017) x US-09-291-922-24 (1-523)

```
Qy          99 GCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCC 158
               |||          |||   :::          |||||||||:::          ::::::|||
Db          12 AlaHisLysThrLeuGlnAspPheAspProProLysLysArgLysArgAsnLysTyrAla 31

Qy         159 TTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTG 218
               ||||||||||||:::|||||||||||||||||||||||||||||||||||||||||||||
Db          32 PheAlaCysAlaMetLeuAlaSerMetThrSerIleLeuLeuGlyTyrAspIleGlyVal 51

Qy         219 ATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAG 278
               ||||||||||||:::::||||||||:::|||      :::||||||   ::::::|||
Db          52 MetSerGlyAlaAlaIleTyrIleLysArgAspLeuLysValSerAspGluGlnIleGlu 71

Qy         279 GTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACG 338
               :::|||:::||||||:::||||||||||||||||||||      ||||||||||||
Db          72 IleLeuLeuGlyIleIleAsnLeuTyrSerLeuIleGlySerCysLeuAlaGlyArgThr 91

Qy         339 TCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGG 398
               ||||||||||||||   ||||||||||||||||||||      |||||||||   |||
Db          92 SerAspTrpIleGlyProArgTyrThrIleValPheAlaGlyThrIlePhePheValGly 111

Qy         399 GSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCC 458
                     ||||||||||:::   ||||||:::   ||||||||||||||||||||||
Db         112 AlaLeuLeuMetGlyPheSerProAsnTyrSerPheLeuMetPheGlyArgPheValAla 131

Qy         459 GGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCG 518
               |||||||||:::||||||||||||||||||||||||||||||||||||||||||||||||
Db         132 GlyIleGlyIleGlyTyrAlaLeuMetIleAlaProValTyrThrAlaGluValSerPro 151
```

```
Qy        519 GCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTG 578
Db        152 AlaSerSerArgGlyPheLeuThrSerPheProGluValPheIleAsnGlyGlyIleLeu 171

Qy        579 CTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATC 638
Db        172 IleGlyTyrIleSerAsnTyrAlaPheSerLysLeuThrLeuLysValGlyTrpArgMet 191

Qy        639 ATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCG 698
Db        192 MetLeuGlyValGlyAlaIleProSerValLeuLeuThrValGlyValLeuAlaMetPro 211

Qy        699 GAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAG 758
Db        212 GluSerProArgTrpLeuValMetArgGlyArgLeuGlyGluAlaArgLysValLeuAsn 231

Qy        759 AAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCC 818
Db        232 LysThrSerAspSerLysGluGluAlaGlnLeuArgLeuAlaGluIleLysGlnAlaAla 251

Qy        819 GGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAAC 878
Db        252 GlyIleProGluSerCysAsnAspAspValValGlnValAsnLysGlnSerAsnGly--- 270

Qy        879 GAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTG 938
Db        271 ---GluGlyValTrpLysGluLeuPheLeuTyrProThrProAlaIleArgHisIleVal 289

Qy        939 CTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTC 998
Db        290 IleAlaAlaLeuGlyIleHisPhePheGlnGlnAlaSerGlyValAspAlaValValLeu 309

Qy        999 TACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGCACC 1058
Db        310 TyrSerProArgIlePheGluLysAlaGlyIleThrAsnAspThrHisLysLeuLeuAla 329

Qy       1059 ACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGAC 1118
Db        330 ThrValAlaValGlyPheValLysThrValPheIleLeuAlaAlaThrPheThrLeuAsp 349

Qy       1119 GGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATC 1178
Db        350 ArgValGlyArgArgProLeuLeuSerSerValGlyGlyMetValLeuSerLeuLeu 369

Qy       1179 GGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCC 1238
Db        370 ThrLeuAlaIleSerLeuThrValIle---AspHisSerGluArgLysLeuMetTrpAla 388

Qy       1239 ATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGC 1298
Db        389 ValGlySerSerIleAlaMetValLeuAlaTyrValAlaThrPheSerIleGlyAlaGly 408

Qy       1299 CCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGC 1358
Db        409 ProIleThrTrpValTyrSerSerGluIlePheProLeuArgLeuArgAlaGlnGlyAla 428

Qy       1359 TCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCG 1418
Db        429 AlaAlaGlyValAlaValAsnArgThrThrSerAlaValValSerMetThrPheLeuSer 448

Qy       1419 CTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTC 1478
Db        449 LeuThrArgAlaIleThrIleGlyGlyAlaPhePheLeuTyrCysGlyIleAlaThrVal 468

Qy       1479 GCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATG 1538
Db        469 GlyTrpIlePhePheTyrThrValLeuProGluThrArgGlyLysThrLeuGluAspMet 488

Qy       1539 AGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAGCCAAGGAGAAG 1598
Db        489 GluGlySerPheGlyThrPheArgSerLysSerAsnAla------------------Ser 502

Qy       1599 AAGAAGGTGGAAATGGCCGCCACTAACTGATCAAACTAACCGCAAAATCACCAAATCCTA 1658
```

```
Db        503 LysAlaValGlu--------------------------AsnGluAsnGlyGlnValAla 513

Qy       1659 AGGGTTTTCTTGCAAAAACGTGTG 1682
Db        514 GlnValGlnLeuGlyThrAsnVal 521
```

RESULT 6
US-09-291-922-30
; Sequence 30, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 30
; LENGTH: 549
; TYPE: PRT
; ORGANISM: Beta vulgaris
US-09-291-922-30

Alignment Scores:

Pred. No.:	6.89e-114	Length:	549
Score:	1589.00	Matches:	326
Percent Similarity:	72.23%	Conservative:	59
Best Local Similarity:	61.16%	Mismatches:	126
Query Match:	43.97%	Indels:	22
DB:	4	Gaps:	5

US-10-051-902A-21 (1-2017) x US-09-291-922-30 (1-549)

```
Qy         78 TCCTCTGCACCACCGGAGATGGCTTCCGCCGCGCTG-----------CCGGAGGCCGTC 125
Db         10 SerAspProProProThrThrAlaSerLysValIleAlaAspPheAspProLeuLysLys 29

Qy        126 GCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATG 185
Db         30 ProProLysArgAsn---------LysPheAlaPheAlaCysAlaThrLeuAlaSerMet 46

Qy        186 ACCTCCATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAG 245
Db         47 ThrSerValLeuLeuGlyTyrAspIleGlyValMetSerGlyAlaIleIleTyrLeuLys 66

Qy        246 AAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTAC 305
Db         67 GluAspTrpHisIleSerAspThrGlnIleGlyValLeuValGlyIleLeuAsnIleTyr 86

Qy        306 TCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACC 365
Db         87 CysLeuPheGlySerPheAlaAlaGlyArgThrSerAspTrpIleGlyArgArgTyrThr 106

Qy        366 ATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAAC 425
Db        107 IleValLeuAlaGlyAlaIlePhePheValGlyAlaLeuLeuMetGlyPheAlaThrAsn 126

Qy        426 TACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATG 485
Db        127 TyrAlaPheLeuMetValGlyArgPheValThrGlyIleGlyValGlyTyrAlaLeuMet 146

Qy        486 ATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCG 545
Db        147 IleAlaProValTyrThrAlaGluValSerProAlaSerSerArgGlyPheLeuThrSer 166

Qy        546 TTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTC 605
```

```
Db        167 PheProGluValPheIleAsnAlaGlyIleLeuLeuGlyTyrIleSerAsnLeuAlaPhe 186

Qy        606 TCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCC 665
              |||   ||||||   :::|||   ||||||   ||||||||||||||||||   ||||||
Db        187 SerSerLeuProThrHisLeuSerTrpArgPheMetLeuGlyIleGlyAlaIleProSer 206

Qy        666 GTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAG 725
              :::   ||||||:::   ||||||   ||||||||||||||||||||||||||||:::
Db        207 IlePheLeuAlaIleGlyValLeuAlaMetProGluSerProArgTrpLeuValMetGln 226

Qy        726 GGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCC 785
              |||||||||   |||||||||   ||||||   :::   ||||||:::   |||||||||
Db        227 GlyArgLeuGlyAspAlaLysLysValLeuAsnArgIleSerAspSerProGluGluAla 246

Qy        786 GCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGAC 845
              |||||||::::::||||||      ||||||||||||   |||   |||   |||
Db        247 GlnLeuArgLeuSerGluIleLysGlnThrAlaGlyIleProAlaGluCysAspGluAsp 266

Qy        846 GTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATC 905
              :::      |||   |||           :::      ||||||||||||||
Db        267 IleTyrLysValGluLysThrLysIleLysSerGlyAsnAlaValTrpLysGluLeuPhe 286

Qy        906 CTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTC 965
                 :::|||||||||||||:::||||||   :::::::::||||||||||||||||||||
Db        287 PheAsnProThrProAlaValArgArgAlaValIleAlaGlyIleGlyIleHisPhePhe 306

Qy        966 CAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCC 1025
              |||   |||   ||||||   :::||||||   |||||||||   :::|||:::|||
Db        307 GlnGlnAlaSerGlyIleAspAlaValValLeuTyrSerProArgIlePheGlnSerAla 326

Qy       1026 GGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGG 1085
              |||:::||||||   :::      |||      |||          ||||||   |||
Db        327 GlyIleThrAsnAlaArgLysGlnLeuLeuAlaThrValAlaValGlyValValLysThr 346

Qy       1086 CTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTG 1145
              ||||||||||:::|||||||||   :::|||      ||||||||||||||||||||
Db        347 LeuPheIleLeuValAlaThrPheGlnLeuAspLysTyrGlyArgArgProLeuLeuLeu 366

Qy       1146 GGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTC 1205
                 |||   ||||||:::|||:::::::::::   |||          |||||||||:::
Db        367 ThrSerValGlyGlyMetIleIleAlaIleLeuThrLeuAlaMetSerLeuThrValIle 386

Qy       1206 GGCCAGCACCCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTC 1265
                    |||          ||||||   |||||||||   |||   |||
Db        387 ---AspHisSerHisHisLysIleThrTrpAlaIleAlaLeuCysIleThrMetValCys 405

Qy       1266 GCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAG 1325
              |||   ||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db        406 AlaValValAlaSerPheSerIleGlyLeuGlyProIleThrTrpValTyrSerSerGlu 425

Qy       1326 ATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTC 1385
              :::|||||||||:::::||||||   |||   |||:::||||||||||   |||||||||
Db        426 ValPheProLeuArgLeuArgAlaGlnGlyThrSerMetGlyValAlaValAsnArgVal 445

Qy       1386 ACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGC 1445
                 ||||||||||||||:::   ||||||   ||||||      ||||||   ||||||
Db        446 ValSerGlyValIleSerIlePhePheLeuProLeuSerHisLysIleThrThrGlyGly 465

Qy       1446 AGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTC 1505
              :::|||||||||:::   |||||||||   :::||||||   ||||||   |||:::|||
Db        466 AlaPhePheLeuPheGlyGlyIleAlaIleIleAlaTrpPhePhePheLeuThrPheLeu 485

Qy       1506 CCGGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGAC--------- 1556
              |||||||||||||||||||||||||   |||   :::||||||   |||
Db        486 ProGluThrArgGlyArgThrLeuGluAsnMetHisGluLeuPheGluAspPheArgTrp 505

Qy       1557 ------------ACGGCCGCCGCCTCGGAATCAGACGAGCCAGCC--------------- 1589
                                  |||   :::||||||   :::
Db        506 ArgGluSerPheProGlyAsnLysSerAsnAsnAspGluAsnSerThrArgLysGlnSer 525

Qy       1590 ------AAGGAGAAGAAGAAGGTGGAAATGGCCGCCACT 1622
                    :::|||   :::|||::::::      |||
```

```
Db        526 AsnGlyAsnAspLysSerGlnValGlnLeuGlyGluThr 538

RESULT 7
US-09-291-922-18
; Sequence 18, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
;  APPLICANT: Allen, Steve
;  APPLICANT: Hitz, Bill
;  APPLICANT: Kinney, Tony
;  APPLICANT: Tingey, Scott
;  TITLE OF INVENTION: Plant Sugar Transport Proteins
;  FILE REFERENCE: BB-1163
;  CURRENT APPLICATION NUMBER: US/09/291,922
;  CURRENT FILING DATE: 1999-04-14
;  EARLIER APPLICATION NUMBER: 60/083,044
;  EARLIER FILING DATE: April 24, 1998
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 18
;   LENGTH: 167
;   TYPE: PRT
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: (34)
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: (85)
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: (98)
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: (112)
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: (151)
US-09-291-922-18

Alignment Scores:
Pred. No.:                1.93e-44        Length:          167
Score:                    682.00          Matches:         138
Percent Similarity:       88.34%          Conservative:    6
Best Local Similarity:    84.66%          Mismatches:      19
Query Match:              18.87%          Indels:          0
DB:                       4               Gaps:            0

US-10-051-902A-21 (1-2017) x US-09-291-922-18 (1-167)

Qy         75 TCCTCCTCTGCACCACCGGAGATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAG 134
              :::   |||   |||      ||||||||||   ||||||   |||:::   |||
Db          3 AlaGlnSerGluProSerThrMetAlaSerAlaProLeuProAlaAlaIleGluProGly 22

Qy        135 AAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATC 194
              ||||||||||||||:::|||||||||||||||   ||||||||||||||||||||||||
Db         23 LysLysGlyAsnValLysPheAlaPheAlaCys***IleLeuAlaSerMetThrSerIle 42

Qy        195 CTCCTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTC 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         43 LeuLeuGlyTyrAspIleGlyValMetSerGlyAlaSerLeuTyrIleLysLysAspLeu 62

Qy        255 AACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATC 314
                 |||||||||   |||:::|||:::|||||||||||||||||:::||||||||||||
Db         63 LysIleSerAspValLysLeuGluIleLeuMetGlyIleLeuAsnValTyrSerLeuIle 82

Qy        315 GGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTC 374
              ||||||   ||||||||||||||||||||||||||||||||   ||||||||||||
Db         83 GlySer***AlaAlaGlyArgThrSerAspTrpIleGlyArgArg***ThrIleValPhe 102

Qy        375 GCCGCCGTCATATTCTTCGCGGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATG 434
              ||||||||||||||||||||||      ||||||||||||||||||||||||   |||
Db        103 AlaAlaValIlePhePheAlaGlyAla***LeuMetGlyPheAlaValAsnTyrTrpMet 122
```

```
Qy        435 CTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCG 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 LeuMetPheGlyArgPheValAlaGlyIleGlyValGlyTyrAlaLeuMetIleAlaThr 142

Qy        495 GTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAG 554
              ||||||||||||||||||||||   ||||||||||||||||||||||||||||||||||
Db        143 ValTyrThrAlaGluValSerPro***SerAlaArgGlyPheLeuThrSerPheProGlu 162

Qy        555 GTGTTCATC 563
              |||||||||
Db        163 ValPheIle 165

RESULT 8
US-10-162-012-46
; Sequence 46, Application US/10162012
; Patent No. 6682597
; GENERAL INFORMATION:
;  APPLICANT: Curtis, Rory A.J.
;  APPLICANT:  Silos-Santiago, Inmaculada
;  APPLICANT:  Gu, Wei
;  TITLE OF INVENTION: NOVEL HUMAN ION CHANNEL AND TRANSPORTER FAMILY MEMBERS
;  FILE REFERENCE: 10448-190001
;  CURRENT APPLICATION NUMBER: US/10/162,012
;  CURRENT FILING DATE:  2002-06-04
;  PRIOR APPLICATION NUMBER: US 60/209,845
;  PRIOR FILING DATE: 2000-06-06
;  PRIOR APPLICATION NUMBER: US 09/875,321
;  PRIOR FILING DATE: 2001-06-06
;  PRIOR APPLICATION NUMBER: PCT/US01/18340
;  PRIOR FILING DATE: 2001-06-06
;  PRIOR APPLICATION NUMBER: US 60/209,257
;  PRIOR FILING DATE: 2000-06-05
;  PRIOR APPLICATION NUMBER: US 09/875,423
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: PCT/US01/18398
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US 60/209,238
;  PRIOR FILING DATE: 2000-06-05
;  PRIOR APPLICATION NUMBER: US 09/875,363
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: PCT/US01/18247
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US 60/227,068
;  PRIOR FILING DATE: 2000-08-22
;  PRIOR APPLICATION NUMBER: US 09/928,530
;  PRIOR FILING DATE: 2001-08-13
;  PRIOR APPLICATION NUMBER: PCT/US01/25475
;  PRIOR FILING DATE: 2001-08-15
;  PRIOR APPLICATION NUMBER: US 60/226,770
;  PRIOR FILING DATE: 2000-08-21
;  PRIOR APPLICATION NUMBER: US 09/934,421
;  PRIOR FILING DATE: 2001-08-21
;  PRIOR APPLICATION NUMBER: PCT/US01/26096
;  PRIOR FILING DATE: 2001-08-21
;  PRIOR APPLICATION NUMBER: US 60/279,281
;  PRIOR FILING DATE: 2001-03-28
;  PRIOR APPLICATION NUMBER: US 10/109,029
;  PRIOR FILING DATE: 2002-03-28
;  PRIOR APPLICATION NUMBER: PCT/US02/09728
;  PRIOR FILING DATE: 2002-03-28
;  PRIOR APPLICATION NUMBER: US 60/290,288
;  PRIOR FILING DATE: 2001-05-11
;  PRIOR APPLICATION NUMBER: US (not assigned)
;  PRIOR FILING DATE: 2002-05-13
;  NUMBER OF SEQ ID NOS: 48
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 46
;   LENGTH: 488
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: consensus sequence

US-10-162-012-46

Alignment Scores:
Pred. No.:                1.69e-43      Length:          488
Score:                    672.00        Matches:         178
Percent Similarity:       53.98%        Conservative:    93
Best Local Similarity:    35.46%        Mismatches:      175
Query Match:              18.59%        Indels:          56
DB:                       4             Gaps:            17

US-10-051-902A-21 (1-2017) x US-10-162-012-46 (1-488)

Qy        168 GCCATCCTCGCCTCCATG---ACCTCCATCCTCCTCGGCTACGATATCGGGGTGATGAGC 224
              |||::::::|||::::::         |||   |||||||||   ||||||:::
Db          2 AlaLeuValAlaAlaLeuGlyGlyGlyPheLeuPheGlyTyrAspThrGlyValIleGly 21

Qy        225 GGGGCGTCGCTGTACATCAAGAAGGACTTCAACATC-----------AGTGACGGGAAG 272
              |||           |||         |||                 |||   |||
Db         22 GlyPheLeuAlaLeuIleAspPheLeuPheArgPheGlyLeuLeuThrSerSerGlyAla 41

Qy        273 GTGGAGGTTCTCATGGGC---------------------------------ATACTGAAC 299
              :::     |||:::|||                                 |||
Db         42 LeuAlaGluLeuValGlyTyrSerThrValLeuThrGlyLeuValValSerIlePhePhe 61

Qy        300 CTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGG 359
              |||      ||||||||||||      ||||||:::      |||      ||||||:::
Db         62 LeuGlyArgLeuIleGlySerLeuPheAlaGlyLysLeuGlyAspArgPheGlyArgLys 81

Qy        360 TACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCC 419
                 :::::::::   |||   |||:::|||      |||      |||   |||   |||
Db         82 LysSerLeuLeuIleAlaLeuValLeuPheValIleGlyAlaLeuLeuSerGlyAlaAla 101

Qy        420 GTCAACTACGCC----------------------ATGCTCATGTTCGGCCGCTTCGTG 455
                  |||                          :::|||:::   ||||||   :::
Db        102 ProGlyTyrThrThrIleGlyLeuTrpAlaPheTyrLeuLeuIleValGlyArgValLeu 121

Qy        456 GCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCG 515
                 |||:::|||||||||   |||   ::::::   |||:::|||   :::|||::::::
Db        122 ValGlyLeuGlyValGlyGlyAlaSerValLeuValProMetTyrIleSerGluIleAla 141

Qy        516 CCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATC 575
              |||   :::   ||||||   |||   |||      ::::::   |||      ||||||
Db        142 ProLysAlaLeuArgGlyAlaLeuGlySerLeuTyrGlnLeuAlaIleThrIleGlyIle 161

Qy        576 CTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTG-----------CCGCTGAAC 623
              |||:::      :::                 ::::::                ||||||
Db        162 LeuValAlaAlaIleIleGlyLeuGlyLeuAsnLysThrAsnAsnAspSerAlaLeuAsn 181

Qy        624 ---CTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTC 680
                  |||||||||||   ||||||:::        |||::::::||||||   :::
Db        182 SerTrpGlyTrpArgIleProLeuGlyLeuGlnLeuValProAlaLeuLeuLeuLeuIle 201

Qy        681 ATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGAC 740
              :::|||   :::||||||||||||||||||||||||   ||||||:::|||   :::
Db        202 GlyLeuLeuPheLeuProGluSerProArgTrpLeuValGluLysGlyLysLeuGluGlu 221

Qy        741 GCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCC 800
              |||:::   ||||||   |||                |||:::      :::   :::
Db        222 AlaArgGluValLeuAlaLysLeuArgGly---ValGluAspValAspGlnGluIleGln 240

Qy        801 GACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCC 860
              :::|||||||||                     ||||||:::                ||||||
Db        241 GluIleLysAla-----------------GluLeuGlu---------AlaThrValSer 251

Qy        861 AAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACC--- 917
              :::      :::|||      |||      |||   ||||||                |||
Db        252 GluGluLysAlaGly------LysAlaSerTrpGlyGluLeuPheArgGlyArgThrArg 269

Qy        918 CCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTG 977
              |||   :::|||:::   ||||||   |||:::   :::      ||||||
Db        270 ProLysValArgGlnArgLeuLeuMetGlyValMetLeuGlnAlaPheGlnGlnLeuThr 289

Qy        978 GGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAAC 1037
```

```
Db        290 GlyIleAsnAlaIlePheTyrTyrSerProThrIlePheLysSerValGlyValSerAsp 309

Qy       1038 GACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTG 1097
Db        310 SerValAlaSerLeuLeuValThrIleIleValGlyValValAsnPheValPheThrPhe 329

Qy       1098 TTGGCG---ACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACG 1154
Db        330 ValAlaLeuIlePheLeuValAspArgPheGlyArgArgProLeuLeuLeuGlyAla 349

Qy       1155 GGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTC--------- 1205
Db        350 AlaGlyMetAlaIleCysPheLeuIleLeuGlyAlaSerIleGlyValAlaLeuLeuLeu 369

Qy       1206 ---------GGCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTA---AGCATC 1253
Db        370 LeuAsnLysProLysAspProSerSerLys---------AlaAlaGlyIleValAlaIle 386

Qy       1254 GCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTG 1313
Db        387 ValPheIleLeuLeuPheIleAlaPhePheAlaLeuGlyTrpGlyProIleProTrpVal 406

Qy       1314 TACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCC 1373
Db        407 IleLeuSerGluLeuPheProThrLysValArgSerLysAlaLeuAlaLeuAlaThrAla 426

Qy       1374 GCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATC 1433
Db        427 AlaAsnTrpLeuAlaAsnPheIleIleGlyPheLeuPheProTyrIleThrGlyAlaIle 446

Qy       1434 ------ACCATCGGCGGCAGCTTCTTCCTC---TACTCCGGCATCGCCGCGCTCGCCTGG 1484
Db        447 GlyLeuAlaLeuGlyGlyTyrValPheLeuValPheAlaGlyLeuLeuValLeuPheIle 466

Qy       1485 GTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAG 1544
Db        467 LeuPheValPhePhePheValProGluThrLysGlyArgThrLeuGluGluIleGluGlu 486

Qy       1545 CTGTTC 1550
Db        487 LeuPhe 488
```

RESULT 9
US-09-489-039A-11902
; Sequence 11902, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 11902
; LENGTH: 514
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11902

Alignment Scores:

Pred. No.:	2.85e-37	Length:	514
Score:	591.00	Matches:	147
Percent Similarity:	51.82%	Conservative:	109
Best Local Similarity:	29.76%	Mismatches:	194
Query Match:	16.35%	Indels:	44
DB:	4	Gaps:	11

US-10-051-902A-21 (1-2017) x US-09-489-039A-11902 (1-514)

```
Qy         93 GAGATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAG---AAGAAGGGCAACGTC 149
Db         41 LysMetThrSerIleSerAsnAspSerThrLeuSerProArgThrGlnArgAspThrArg 60

Qy        150 CGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGAT 209
Db         61 ArgMetAsnTrpPheValSerIleAlaAlaAlaValAlaGlyLeuLeuPheGlyLeuAsp 80

Qy        210 ATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGG 269
Db         81 IleGlyValIleSerGlyAlaLeuProPheIleThrAspHisPheThrLeuSerSerGln 100

Qy        270 AAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCG 329
Db        101 LeuGlnGluTrpValValSerSerMetMetLeuGlyAlaAlaIleGlyAlaLeuPheAsn 120

Qy        330 GGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTC 389
Db        121 GlyTrpLeuSerPheArgLeuGlyArgLysTyrSerLeuMetAlaGlyAlaValLeuPhe 140

Qy        390 TTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGC 449
Db        141 ValAlaGlySerIleGlySerAlaPheAlaAlaSerValGluValLeuLeuValAlaArg 160

Qy        450 TTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAG 509
Db        161 ValValLeuGlyValAlaValGlyIleAlaSerTyrThrAlaProLeuTyrLeuSerGlu 180

Qy        510 GTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTC 569
Db        181 MetAlaSerGluAsnValArgGlyLysMetIleSerMetTyrGlnLeuMetValThrLeu 200

Qy        570 GGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGG 629
Db        201 GlyIleValLeuAlaPheLeuSerAspThrAlaPheSer---------TyrSerGlyAsn 217

Qy        630 TGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTC 689
Db        218 TrpArgAlaMetLeuGlyValLeuAlaLeuProAlaValIleLeuIleIleLeuValVal 237

Qy        690 GGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTG 749
Db        238 PheLeuProAsnSerProArgTrpLeuAlaGluLysGlyArgHisIleGluAlaGluGlu 257

Qy        750 GTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAG 809
Db        258 ValLeuArgMetLeuArgAspThrSerGluLysAlaArgAspGluLeuAsnGluIleArg 277

Qy        810 GCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGG 869
Db        278 GluSerLeuLysLeu----------LysGlnGly 285

Qy        870 AGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGG 929
Db        286 Gly----------Trp------AlaLeuPheLysIleAsnArgAsnValArg 297

Qy        930 CGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCC 989
Db        298 ArgAlaValPheLeuGlyMetLeuLeuGlnAlaMetGlnGlnPheThrGlyMetAsnIle 317

Qy        990 GTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTC 1049
Db        318 IleMetTyrTyrAlaProArgIlePheLysMetAlaGlyPheThrThrThrGluGlnGln 337

Qy       1050 TTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTC 1109
Db        338 MetIleAlaThrLeuValValGlyLeuThrPheMetPheAlaThrPheIleAlaValPhe 357

Qy       1110 TTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTC 1169
Db        358 ThrValAspLysAlaGlyArgLysProAlaLeu------LysIleGlyPheSerValMet 375

Qy       1170 TCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAG---------------- 1211
```

```
Db        376 AlaLeu--------------GlyThrLeuValLeuGlyTyrCysLeuMetGlnPheAsp 390

Qy       1212 ------CACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTC 1265
Db        391 AsnGlyThrAlaSerSerGlyLeuSerTrp---------LeuSerValGlyMetThrMet 407

Qy       1266 GCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAG 1325
Db        408 MetCysIleAlaGlyTyrAlaMetSerAlaAlaProValValTrpIleLeuCysSerGlu 427

Qy       1326 ATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTC 1385
Db        428 IleGlnProLeuLysCysArgAspPheGlyIleThrCysSerThrThrThrAsnTrpVal 447

Qy       1386 ACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGC 1445
Db        448 SerAsnMetIleIleGlyAlaThrPheLeuThrLeuLeuAspAlaIleGlyAlaAlaGly 467

Qy       1446 AGCTTCTTCCTCTACTCCGGCATC---GCCGCGCTCGCCTGGGTGTTCTTCTACACCTAC 1502
Db        468 ThrPheTrpLeuTyrThrAlaLeuAsnValAlaPheIleGlyIleThrPheTrp---Leu 486

Qy       1503 CTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAG 1544
Db        487 IleProGluThrLysAsnValThrLeuGluHisIleGluArg 500
```

RESULT 10
US-09-489-039A-11933
; Sequence 11933, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 11933
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11933

Alignment Scores:

Pred. No.:	7.99e-37	Length:	476
Score:	585.00	Matches:	146
Percent Similarity:	49.49%	Conservative:	97
Best Local Similarity:	29.74%	Mismatches:	212
Query Match:	16.19%	Indels:	36
DB:	4	Gaps:	8

US-10-051-902A-21 (1-2017) x US-09-489-039A-11933 (1-476)

```
Qy         90 CCGGAGATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCG------AAGAAGAAGGGC 143
Db          1 ProSerThrGluSerIleThrGlnLeuGluGlyValMetProAspAsnLysLysGlnGly 20

Qy        144 ------AACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTC 197
Db         21 ArgSerAsnLysThrMetThrPhePheValCysPheLeuAlaAlaLeuAlaGlyLeuLeu 40

Qy        198 CTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAAC 257
Db         41 PheGlyLeuAspIleGlyValIleAlaGlyAlaLeuProPheIleAlaAsnGluPheGln 60

Qy        258 ATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGC 317
Db         61 IleSerAlaHisThrGlnGluTrpValValSerSerMetMetPheGlyAlaAlaValGly 80

Qy        318 TCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCC 377
Db         81 AlaValGlySerGlyTrpLeuSerPheLysLeuGlyArgLysLysSerLeuMetIleGly 100

Qy        378 GCCGTCATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTC 437
Db        101 AlaIleLeuPheValAlaGlySerLeuPheSerAlaAlaAlaProAsnValGluIleLeu 120

Qy        438 ATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTG 497
Db        121 LeuValSerArgValLeuLeuGlyLeuAlaValGlyValAlaSerTyrThrAlaProLeu 140

Qy        498 TACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTG 557
Db        141 TyrLeuSerGluIleAlaProGluLysIleArgGlySerMetIleSerMetTyrGlnLeu 160

Qy        558 TTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCG 617
Db        161 MetIleThrIleGlyIleLeuGlyAlaTyrLeuSerAspThrAlaPheSer--------- 177

Qy        618 CTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCG 677
Db        178 TyrSerGlyAlaTrpArgTrpMetLeuGlyValIleIleIleProAlaValLeuLeuLeu 197

Qy        678 CTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCG 737
Db        198 IleGlyValIlePheLeuProAspSerProArgTrpPheAlaAlaLysArgArgPheVal 217

Qy        738 GACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTG 797
Db        218 AspAlaGluArgValLeuLeuArgLeuArgAspThrSerAlaGluAlaLysArgGluLeu 237

Qy        798 GCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTC 857
Db        238 AspGluIleArgGluSerLeuLysValLysGln---------------------------- 248

Qy        858 CCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACC 917
Db        249 -----------SerGly--------------TrpSer------LeuPheLysAspAsn 257

Qy        918 CCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTG 977
Db        258 SerAsnPheArgArgAlaValPheLeuGlyIleLeuLeuGlnValMetGlnGlnPheThr 277

Qy        978 GGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAAC 1037
Db        278 GlyMetAsnValIleMetTyrTyrAlaProLysIlePheGluLeuAlaGlyTyrAlaAsn 297

Qy       1038 GACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTG 1097
Db        298 ThrThrGluGlnMetTrpGlyThrValIleValGlyLeuThrAsnValLeuAlaThrPhe 317

Qy       1098 TTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGC 1157
Db        318 IleAlaIleGlyLeuValAspArgTrpGlyArgLysProThrLeu--------------- 332

Qy       1158 GGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGC------CAG 1211
Db        333 ------IleLeuGlyPheIleValMetAlaAlaGlyMetGlyValLeuGlyThrMetMet 350

Qy       1212 CACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTAC 1271
Db        351 HisIleGlyIleHisSerSerThrAlaGlnTyrIleAlaValLeuMetLeuLeuMetPhe 370

Qy       1272 GTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTC 1331
Db        371 IleValGlyPheAlaMetSerAlaGlyProLeuIleTrpValLeuCysSerGluIleGln 390

Qy       1332 CCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGC 1391
Db        391 ProLeuLysGlyArgAspPheGlyIleThrCysSerThrAlaThrAsnTrpIleAlaAsn 410

Qy       1392 GGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTC 1451
```

```
                    ::::::      ||||||||::::::      ::::::            :::|||
Db        411 MetIleValGlyAlaThrPheLeuThrMetLeuAsnSerLeuGlySerAlaAsnThrPhe 430

Qy       1452 TTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAG 1511
                ::::::|||   |||:::      |||      :::             :::||||||
Db        431 TrpValTyrGlyGlyLeuAsnValLeuPheIleLeuLeuThrLeuTrpLeuIleProGlu 450

Qy       1512 ACCCGCGGCCGGACGCTGGAGGAGATGAGCAAG 1544
                |||:::      :::||||||   :::   :::
Db        451 ThrLysAsnValSerLeuGluHisIleGluArg 461

RESULT 11
US-08-928-692-13
; Sequence 13, Application US/08928692
; Patent No. 5958727
;  GENERAL INFORMATION:
;    APPLICANT:  Brody, Howard
;    APPLICANT:  Yaver, Deborah S.
;    APPLICANT:  Lamsa, Michael
;    APPLICANT:  Hansen, Kim
;    TITLE OF INVENTION:  Methods for Modifying the Production of
;    TITLE OF INVENTION:  a Polypeptide
;    NUMBER OF SEQUENCES:  80
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 5958727c No. 5958727disk of No. 5958727th America, Inc.
;      STREET:  405 Lexington Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10174
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/928,692
;      FILING DATE:  12-SEPT-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lambiris, Elias J
;      REGISTRATION NUMBER:  33,728
;      REFERENCE/DOCKET NUMBER:  4944.200-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-867-0123
;      TELEFAX:  212-878-9655
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  584 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 5958727e
US-08-928-692-13

Alignment Scores:
Pred. No.:                4.23e-36      Length:          584
Score:                    576.00        Matches:         148
Percent Similarity:       50.40%        Conservative:    106
Best Local Similarity:    29.37%        Mismatches:      206
Query Match:              15.94%        Indels:          44
DB:                       2             Gaps:            11

US-10-051-902A-21 (1-2017) x US-08-928-692-13 (1-584)

Qy        153 TTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATC 212
                |||                 :::||||||::::::      :::   |||||||||
Db         85 PheIleIleThrLeuThrPheValAlaSerIleSerGlyPheMetPheGlyTyrAspThr 104

Qy        213 GGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAAC---------ATCAGT 263
                |||   :::|||   |||   :::   |||       |||   :::          ::::::
Db        105 GlyTyrIleSerSerAlaLeuIleSerIleGlyThrAspLeuAspHisLysValLeuThr 124

Qy        264 GACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTC 323
                |||:::   |||::::::              :::|||   :::||||||   |||
Db        125 TyrGlyGluLysGluIleValThrAlaAlaThrSerLeuGlyAlaLeuIleThrSerIle 144

Qy        324 GCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTC 383
                ||||||      :::|||      ||||||:::      ::::::   :::   :::
Db        145 PheAlaGlyThrAlaAlaAspIlePheGlyArgLysArgCysLeuMetGlySerAsnLeu 164

Qy        384 ATATTCTTCGCGGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTC 443
                :::|||      |||      |||      |||      :::      :::
Db        165 MetPheValIleGlyAlaIleLeuGlnValSerAlaHisThrPheTrpGlnMetAlaVal 184

Qy        444 GGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACC 503
                ||||||   :::   |||   ||||||||         :::|||||||||::::::
Db        185 GlyArgLeuIleMetGlyPheGlyValGlyIleGlySerLeuIleAlaProLeuPheIle 204

Qy        504 GCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATC 563
                :::|||::::::|||         ||||||   ||||||                :::::::::
Db        205 SerGluIleAlaProLysMetIleArgGlyArgLeuThrValIleAsnSerLeuTrpLeu 224

Qy        564 AACTTCGGCATCCTGCTCGGGTACGTCTCG-----------AACTATGCTTTCTCCCGC 611
                  |||   |||:::   |||                   ||||||
Db        225 ThrGlyGlyGlnLeuValAlaTyrGlyCysGlyAlaGlyLeuAsnTyr----------- 240

Qy        612 TTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTG 671
                  :::|||   ||||||||||:::::|||:::         |||:::   :::
Db        241 ------ValAsnAsnGlyTrpArgIleLeuValGlyLeuSerLeuIleProThrAlaVal 258

Qy        672 CTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGC 731
                        :::      :::|||:::::||||:::   ||||||||||
Db        259 GlnPheThrCysLeuCysPheLeuProAspThrProArgTyrTyrValMetLysGlyAsp 278

Qy        732 CTCGCGGACGCCAAGGTGGTGCTGGAGAAGACC---TCCGACACGGCGGAGGAGGCCGCG 788
                ||||||   |||      ||||||:::::::::   :::||||||:::||||||
Db        279 LeuAlaArgAlaThrGluValLeuLysArgSerTyrThrAspThrSerGluGluIleIle 298

Qy        789 GAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTG 848
                ||||||      :::                                      ::::::
Db        299 GluArgLysValGlu----------------------------------------GluLeu 305

Qy        849 GTGACCGTCCCCAAGAGA------GGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTC 902
                ||||||:::   :::         |||           ::::::||||||      :::
Db        306 ValThrLeuAsnGlnSerIleProGlyLysAsnValProGluLysValTrpAsnThrIle 325

Qy        903 ---ATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCAC 959
                   |||         |||:::      |||   |||:::   |||   |||:::
Db        326 LysGluLeuHisThrValProSerAsnLeuArgAlaLeuIleIleGlyCysGlyLeuGln 345

Qy        960 TTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAG 1019
                   |||         |||   :::|||:::::::::::::|||      :::|||:::
Db        346 AlaIleGlnGlnPheThrGlyTrpAsnSerLeuMetTyrPheSerGlyThrIlePheGlu 365

Qy       1020 AGCCCCGGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACC 1079
                :::   |||      |||                        :::          |||
Db        366 ThrValGlyPheLysAsnSerSer---------AlaValSerIleIleValSerGlyThr 382

Qy       1080 AAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTG 1139
                   :::|||   |||:::|||   |||   ||||||   :::|||||||||   :::
Db        383 AsnPheIlePheThrLeuValAlaPhePheSerIleAspLysIleGlyArgArgThrIle 402

Qy       1140 TTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACC 1199
                ||||||            |||:::   ::::::|||:::
Db        403 LeuLeuIleGlyLeuProGlyMetThrMetAlaLeuValValCysSerIleAlaPheHis 422

Qy       1200 GTCGTCGGCCAGCACCCCGACGCCAAGATA--------------------------CCT 1232
                :::|||         |||      :::.
Db        423 PheLeuGlyIleLysPheAspGlyAlaValAlaValValValSerSerGlyPheSerSer 442

Qy       1233 TGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGC 1292
                |||   |||   :::   |||          :::   :::   ||||||::::::::|||
Db        443 TrpGlyIle---ValIleIleValPheIleIleValPheAlaAlaPheTyrAlaLeuGly 461

Qy       1293 CTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTG 1352
```

```
                :::|||   :::  |||         ||||||:::||||||      ||||||   :::
Db        462 IleGlyThrValProTrp---GlnGlnSerGluLeuPheProGlnAsnValArgGlyIle 480

Qy       1353 GGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTC 1412
              |||   |||          |||   |||          |||   ||||||:::   ||||||
Db        481 GlyThrSerTyrAlaThrAlaThrAsnTrpAlaGlySerLeuValIleAlaSerThrPhe 500

Qy       1413 CTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCC 1472
              |||::::::   :::   ||||||      |||:::|||      ::::::|||::::::
Db        501 LeuThrMetLeuGlnAsnIleThrProAlaGlyThrPheAlaPhePheAlaGlyLeuSer 520

Qy       1473 GCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAG 1532
                |||:::   :::|||   |||         ||||||      |||      ||||||
Db        521 CysLeuSerThrIlePheCysTyrPheCysTyrProGluLeuSerGlyLeuGluLeuGlu 540

Qy       1533 GAGATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAGCCAAG 1592
              |||:::      :::      |||                 |||      ||||||
Db        541 GluValGlnThrIleLeuLysAspGlyPheAsnIleLysAlaSerLysAlaLeuAlaLys 560

Qy       1593 GAGAAGAAGAAG 1604
              ::::::|||:::
Db        561 LysArgLysGln 564
```

RESULT 12
US-09-339-972-13
; Sequence 13, Application US/09339972
; Patent No. 6323002
; GENERAL INFORMATION:
; APPLICANT: Brody, Howard
; APPLICANT: Yaver, Deborah S.
; APPLICANT: Lamsa, Michael
; APPLICANT: Hansen, Kim
; TITLE OF INVENTION: Methods for Modifying the Production of
; TITLE OF INVENTION: a Polypeptide
; NUMBER OF SEQUENCES: 80
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6323002o No. 6323002disk of No. 6323002th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/339,972
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/928,692
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4944.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 584 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6323002e
US-09-339-972-13

Alignment Scores:

Pred. No.:	4.23e-36	Length:	584
Score:	576.00	Matches:	148
Percent Similarity:	50.40%	Conservative:	106
Best Local Similarity:	29.37%	Mismatches:	206
Query Match:	15.94%	Indels:	44
DB:	4	Gaps:	11

US-10-051-902A-21 (1-2017) x US-09-339-972-13 (1-584)

```
Qy        153 TTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATC 212
              |||                  :::||||||::::::      :::   |||||||||
Db         85 PheIleIleThrLeuThrPheValAlaSerIleSerGlyPheMetPheGlyTyrAspThr 104

Qy        213 GGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAAC---------ATCAGT 263
              |||   :::|||   |||  :::   |||       |||   :::          ::::::
Db        105 GlyTyrIleSerSerAlaLeuIleSerIleGlyThrAspLeuAspHisLysValLeuThr 124

Qy        264 GACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTC 323
              |||:::   |||::::::          :::|||   :::||||||   |||
Db        125 TyrGlyGluLysGluIleValThrAlaAlaThrSerLeuGlyAlaLeuIleThrSerIle 144

Qy        324 GCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTC 383
              ||||||      :::|||      ||||||:::      ::::::   :::   :::
Db        145 PheAlaGlyThrAlaAlaAspIlePheGlyArgLysArgCysLeuMetGlySerAsnLeu 164

Qy        384 ATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTC 443
              :::|||       |||      |||          |||      :::      :::
Db        165 MetPheValIleGlyAlaIleLeuGlnValSerAlaHisThrPheTrpGlnMetAlaVal 184

Qy        444 GGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACC 503
              ||||||   :::   |||   |||||||||        :::|||||||||:::::::
Db        185 GlyArgLeuIleMetGlyPheGlyValGlyIleGlySerLeuIleAlaProLeuPheIle 204

Qy        504 GCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATC 563
              :::|||::::::|||        ||||||   ||||||                :::::::::
Db        205 SerGluIleAlaProLysMetIleArgGlyArgLeuThrValIleAsnSerLeuTrpLeu 224

Qy        564 AACTTCGGCATCCTGCTCGGGTACGTCTCG-----------AACTATGCTTTCTCCCGC 611
                     |||   |||:::   |||                    ||||||
Db        225 ThrGlyGlyGlnLeuValAlaTyrGlyCysGlyAlaGlyLeuAsnTyr---------- 240

Qy        612 TTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTG 671
                 :::|||   ||||||||||:::::|||:::          |||:::   :::
Db        241 ------ValAsnAsnGlyTrpArgIleLeuValGlyLeuSerLeuIleProThrAlaVal 258

Qy        672 CTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGC 731
                       :::      :::|||::::::||||||:::   ||||||||||
Db        259 GlnPheThrCysLeuCysPheLeuProAspThrProArgTyrTyrValMetLysGlyAsp 278

Qy        732 CTCGCGGACGCCAAGGTGGTGCTGGAGAAGACC---TCCGACACGGCGGAGGAGGCCGCG 788
              ||||||   |||      ||||||:::::::::   :::||||||:::||||||
Db        279 LeuAlaArgAlaThrGluValLeuLysArgSerTyrThrAspThrSerGluGluIleIle 298

Qy        789 GAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTG 848
              ||||||      :::                                      ::::::
Db        299 GluArgLysValGlu----------------------------------------GluLeu 305

Qy        849 GTGACCGTCCCCAAGAGA------GGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTC 902
              ||||||:::   :::        |||          ::::::||||||      :::
Db        306 ValThrLeuAsnGlnSerIleProGlyLysAsnValProGluLysValTrpAsnThrIle 325

Qy        903 ---ATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCAC 959
                    |||         |||:::      |||   |||:::   |||   |||:::
Db        326 LysGluLeuHisThrValProSerAsnLeuArgAlaLeuIleIleGlyCysGlyLeuGln 345

Qy        960 TTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAG 1019
                    |||         |||   :::|||:::::::::::::|||      :::|||:::
Db        346 AlaIleGlnGlnPheThrGlyTrpAsnSerLeuMetTyrPheSerGlyThrIlePheGlu 365

Qy       1020 AGCCCCGGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACC 1079
              :::   |||      |||                  :::                  |||
Db        366 ThrValGlyPheLysAsnSerSer---------AlaValSerIleIleValSerGlyThr 382

Qy       1080 AAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTG 1139
                    :::|||   |||:::|||   |||   ||||||   :::||||||||||   :::
```

```
Db        383 AsnPheIlePheThrLeuValAlaPheSerIleAspLysIleGlyArgArgThrIle 402

Qy       1140 TTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACC 1199
              ||||||           |||:::   ::::::|||:::
Db        403 LeuLeuIleGlyLeuProGlyMetThrMetAlaLeuValValCysSerIleAlaPheHis 422

Qy       1200 GTCGTCGGCCAGCACCCCGACGCCAAGATA-------------------------CCT 1232
                :::|||        |||     :::
Db        423 PheLeuGlyIleLysPheAspGlyAlaValAlaValValValSerSerGlyPheSerSer 442

Qy       1233 TGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGC 1292
              |||   |||   :::   |||         :::   :::   ||||||:::::::::|||
Db        443 TrpGlyIle---ValIleIleValPheIleIleValPheAlaAlaPheTyrAlaLeuGly 461

Qy       1293 CTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTG 1352
              :::|||   :::   |||       ||||||:::||||||      ||||||   :::
Db        462 IleGlyThrValProTrp---GlnGlnSerGluLeuPheProGlnAsnValArgGlyIle 480

Qy       1353 GGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTC 1412
              |||   |||         |||   |||         |||   ||||||:::   ||||||
Db        481 GlyThrSerTyrAlaThrAlaThrAsnTrpAlaGlySerLeuValIleAlaSerThrPhe 500

Qy       1413 CTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCC 1472
              |||::::::   :::   ||||||     |||:::|||      ::::::|||::::::
Db        501 LeuThrMetLeuGlnAsnIleThrProAlaGlyThrPheAlaPhePheAlaGlyLeuSer 520

Qy       1473 GCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAG 1532
                |||:::   :::|||   |||         ||||||      |||     ||||||
Db        521 CysLeuSerThrIlePheCysTyrPheCysTyrProGluLeuSerGlyLeuGluLeuGlu 540

Qy       1533 GAGATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAGCCAAG 1592
              |||:::   :::   |||                 |||     ||||||
Db        541 GluValGlnThrIleLeuLysAspGlyPheAsnIleLysAlaSerLysAlaLeuAlaLys 560

Qy       1593 GAGAAGAAGAAG 1604
              ::::::|||:::
Db        561 LysArgLysGln 564
```

RESULT 13
US-09-679-686B-18
; Sequence 18, Application US/09679686B
; Patent No. 6624343
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen M.
; APPLICANT: Lightner, Jonathan E.
; APPLICANT: Rafalski, J. Antoni
; APPLICANT: Thorpe, Catherine J.
; TITLE OF INVENTION: HEXOSE CARRIER PROTEINS
; FILE REFERENCE: BB1160 US NA
; CURRENT APPLICATION NUMBER: US/09/679,686B
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: 60/081,131
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: PCT/US99/07561
; PRIOR FILING DATE: 1999-04-07
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 18
; LENGTH: 517
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-09-679-686B-18

Alignment Scores:

Pred. No.:	4.79e-34	Length:	517
Score:	549.00	Matches:	144
Percent Similarity:	48.68%	Conservative:	95
Best Local Similarity:	29.33%	Mismatches:	196
Query Match:	15.19%	Indels:	56
DB:	4	Gaps:	11

US-10-051-902A-21 (1-2017) x US-09-679-686B-18 (1-517)

```
Qy        153 TTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATC 212
              |||   |||           ::::::|||:::         ::::::   ||||||||||||
Db         22 PheValPhePheThrCysValValAlaAlaThrGlyGlyLeuIlePheGlyTyrAspIle 41

Qy        213 GGGGTGATGAGCGGGGCGTCGCTG---------TACATCAAGAAGGACTTC---AACATC 260
              |||:::       |||   :::             ::::::||||||   |||     :::
Db         42 GlyIleSerGlyGlyValThrSerMetAsnProPheLeuLysLysPhePheProGluVal 61

Qy        261 AGTGACGGGAAG---------------------------------------------- 272
                 |||   |||
Db         62 TyrAspLysLysGlnMetLysGlySerAlaSerGlnTyrCysLysTyrAspAsnGlnLeu 81

Qy        273 GTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGG 332
              ::::::                 |||   |||   :::|||:::   ||||||   |||
Db         82 LeuGlnThrPheThrSerSerLeuTyrLeuAlaAlaLeuValSerSerPhePheAlaAla 101

Qy        333 CGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTC 392
                 :::      :::|||||:::::::::::          :::   |||
Db        102 ThrValThrArgValValGlyArgLysTrpSerMetPheThrGlyGlyLeuThrPheLeu 121

Qy        393 GCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTC 452
                 |||      |||   |||   |||   |||   |||||||||:::   ||||||
Db        122 IleGlyAlaAlaLeuAsnGlyAlaAlaGluAsnIleAlaMetLeuIleValGlyArgIle 141

Qy        453 GTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTG 512
              :::   |||:::|||||||||:::|||          ||||||||   :::|||:::
Db        142 LeuLeuGlyValGlyValGlyPheAlaAsnGlnSerValProValTyrLeuSerGluMet 161

Qy        513 TCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGC 572
              :::||||||      ||||||   |||         ::::::   |||      |||
Db        162 AlaProAlaArgLeuArgGlyMetLeuAsnIleGlyPheGlnLeuMetIleThrIleGly 181

Qy        573 ATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGG 632
              ||||||        :::   ||||||      :::::::::                ||||||
Db        182 IleLeuAlaAlaAlaLeuIleAsnTyrAspThrAsnLysIleLysAlaGlyTyrGlyTrp 201

Qy        633 CGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGC 692
              ||||||   |||   |||   |||   |||:::   ::::::   |||
Db        202 ArgIleSerLeuAlaIleAlaAlaValProAlaGlyIleIleThrLeuGlySerPhePhe 221

Qy        693 ATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTG 752
              :::|||::::::|||       |||:::   :::|||             |||:::   :::
Db        222 LeuProAspThrProAsnSerLeuIleGluArgGlyHisProGluAlaAlaArgArgMet 241

Qy        753 CTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCC 812
              |||   :::       :::       :::   :::|||      ||||||:::
Db        242 LeuAsnArgIleArgGlySerAspValAspIleSerGluGluTyrAlaAspLeuValVal 261

Qy        813 GCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGC 872
              |||:::                 :::      :::|||     |||
Db        262 AlaSerGlu--------------GluSerLysLeuValGlnHisPro------------ 272

Qy        873 GGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGC 932
                           |||:::   ::::::                          |||
Db        273 -----------------TrpArgAsnIleLeu---------GlnArgLysTyrArgPro 283

Qy        933 ATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTC 992
                 |||         |||   |||   |||||||||         ||||||:::   :::
Db        284 GlnLeuThrMetAlaIleMetIleProPhePheGlnGlnLeuThrGlyIleAsnValIle 303

Qy        993 GTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTG 1052
              :::||||||:::|||::::::|||::::::   |||         |||         :::
Db        304 MetPheTyrAlaProValLeuPheGluThrLeuGlyPheLysGlyAspAlaSerLeuMet 323

Qy       1053 GGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTC 1112
                           |||:::                 |||::::::   |||
Db        324 SerAlaValIleThr---GlyLeuValAsnValPheAlaThrLeuValSerValPheThr 342

Qy       1113 ATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCC 1172
              :::|||   :::|||||||||   |||   |||                ::::::||||||
Db        343 ValAspArgLeuGlyArgArgLysLeuPheLeuGlnGlyGlyThrGlnMetLeuLeuSer 362

Qy       1173 -----------------CTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCAC 1214
```

```
                                 ||||||   :::  :    |||   :::   ||||||:::
Db        363 GlnLeuValValGlyThrLeuIleAlaValLysPheGlyThrSerGlyValGlyGluMet 382

Qy       1215 CCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTC--------- 1265
              |||                        |||   :::   |||      |||
Db        383 ProLys--------------------GlyTyrAlaAlaAlaValValLeuPheIleCys 395

Qy       1266 GCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAG 1325
              |||||||||   |||:::         ||||||:::   |||:::      ||||||
Db        396 LeuTyrValAlaGlyPheAlaTrpSerTrpGlyProLeuGlyTrpLeuValProSerGlu 415

Qy       1326 ATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTC 1385
              ||||||||||:::::|||      |||   |||:::   |||:::   |||   :::
Db        416 IlePheProLeuGluIleArgProAlaGlyGlnSerIleAsnValSerValAsnMetLeu 435

Qy       1386 ACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGC 1445
                 :::   ||||||:::      ||||||:::::          :::         |||
Db        436 PheThrPheValIleAlaGlnAlaPheLeuThrMetLeuCysHisMetLysPhe---Gly 454

Qy       1446 AGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTC 1505
              |||:::   ::::::|||         :::      ||||||          :::|||
Db        455 LeuPheTyrPhePheAlaGlyTrpValValIleMetThrValPheIleAlaLeuPheLeu 474

Qy       1506 CCGGAGACCCGCGGCCGGACGCTGGAGGAGATG 1538
              |||||||||:::         :::|||||||||
Db        475 ProGluThrLysAsnValProIleGluGluMet 485

RESULT 14
US-09-679-686B-22
; Sequence 22, Application US/09679686B
; Patent No. 6624343
; GENERAL INFORMATION:
;  APPLICANT: Allen, Stephen M.
;  APPLICANT:  Lightner, Jonathan E.
;  APPLICANT:  Rafalski, J. Antoni
;  APPLICANT:  Thorpe, Catherine J.
;  TITLE OF INVENTION: HEXOSE CARRIER PROTEINS
;  FILE REFERENCE: BB1160 US NA
;  CURRENT APPLICATION NUMBER: US/09/679,686B
;  CURRENT FILING DATE:  2003-01-16
;  PRIOR APPLICATION NUMBER: 60/081,131
;  PRIOR FILING DATE: 1998-04-09
;  PRIOR APPLICATION NUMBER: PCT/US99/07561
;  PRIOR FILING DATE: 1999-04-07
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 22
;   LENGTH: 514
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
US-09-679-686B-22

Alignment Scores:
Pred. No.:                4.76e-33        Length:          514
Score:                    536.00          Matches:         142
Percent Similarity:       48.79%          Conservative:    99
Best Local Similarity:    28.74%          Mismatches:      193
Query Match:              14.83%          Indels:          60
DB:                       4               Gaps:            14

US-10-051-902A-21 (1-2017) x US-09-679-686B-22 (1-514)

Qy        159 TTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTG 218
              |||             :::   :::         ::::::   |||||||||:::|||:::
Db         24 PheValThrCysPheIleGlyAlaPheGlyGlyLeuIlePheGlyTyrAspLeuGlyIle 43

Qy        219 ATGAGCGGGGCGTCGCTG----------------------------TACATCAAGAAG 248
                    |||   :::                                |||:::   |||
Db         44 SerGlyGlyValThrSerMetGluProPheLeuGluGluPhePheProTyrValTyrLys 63

Qy        249 GACTTCAACATCAGT-----------------------GACGGGAAGGTGGAGGTTCTC 284
                     :::                          |||   ::::::      |||
Db         64 LysMetLysSerAlaHisGluAsnGluTyrCysArgPheAspSerGlnLeuLeuThrLeu 83

Qy        285 ATG---GGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCG 341
                    |||   :::   :::|||:::   |||       |||         :::
Db         84 PheThrSerSerLeuTyrValAlaAlaLeuValSerSerLeuPheAlaSerThrIleThr 103

Qy        342 GACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGSG 401
                  ||||||:::::::::::::                    ||||||   |||
Db        104 ArgValPheGlyArgLysTrpSerMetPheLeuGlyGlyPheThrPhePheIleGlySer 123

Qy        402 TTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGC 461
                  |||||||||   |||   |||||||||:::   ||||||   :::   |||
Db        124 AlaPheAsnGlyPheAlaGlnAsnIleAlaMetLeuLeuIleGlyArgIleLeuLeuGly 143

Qy        462 ATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCG 521
                  |||||||||:::|||           |||||||||   :::|||::::::|||
Db        144 PheGlyValGlyPheAlaAsnGlnSerValProValTyrLeuSerGluMetAlaProPro 163

Qy        522 TCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTC 581
                  :::   ||||||         :::      :::|||   |||   |||||||||::::::
Db        164 AsnLeuArgGlyAlaPheAsnAsnGlyPheGlnValAlaIleIlePheGlyIleValVal 183

Qy        582 GGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATG 641
                  :::   ||||||      :::::::::      |||:::|||||||||||||
Db        184 AlaThrIleIleAsnTyrPheThrAlaGlnMetLysGlyAsnIleGlyTrpArgIleSer 203

Qy        642 CTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAG 701
                  ||||||:::         |||:::|||::::::   :::      |||   :::|||:::
Db        204 LeuGlyLeuAlaCysValProAlaValMetIleMetIleGlyAlaLeuIleLeuProAsp 223

Qy        702 TCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAG 761
                  :::|||      |||:::   :::|||
Db        224 ThrProAsnSerLeuIleGluArgGly-------------------------------- 232

Qy        762 ACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGC 821
                           |||||||||   |||   |||      |||:::
Db        233 ---------TyrThrGluGluAlaLysGluMetLeuGlnSerIleArgGlyThrAsnGlu 249

Qy        822 ATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAG 881
                  :::   ||||||         |||::::::         :::      |||
Db        250 ValAspGluGluPheGln---AspLeuIleAspAlaSerGluGlu---SerLysGlnVal 267

Qy        882 AAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTG 941
                  |||      ||||||   ::::::|||         |||      |||      |||:::
Db        268 LysHisProTrpLysAsnIleMetLeu---------ProArgTyrArgProGlnLeuIle 284

Qy        942 TCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTAC 1001
                           |||   |||||||||         ||||||:::   :::   ||||||
Db        285 MetThrCysPheIleProPhePheGlnGlnLeuThrGlyIleAsnValIleThrPheTyr 304

Qy       1002 AGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGCACCACT 1061
                  :::|||::::::|||::::::   |||         |||      |||   :::
Db        305 AlaProValLeuPheGlnThrLeuGlyPhe---GlySerLysAlaSerLeuLeuSerAla 323

Qy       1062 TGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGC 1121
                        |||:::   :::   |||         ::::::   |||   :::|||
Db        324 MetValThrGlyIleIleGluLeuLeuCysThrPheValSerValPheThrValAspArg 343

Qy       1122 GTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATA------ATCCTCTCCCTC 1175
                  |||||||||   |||   |||         |||||||||      ::::::|||
Db        344 PheGlyArgArgIleLeuPheLeu------GlnGlyGlyIleGlnMetLeuValSerGln 361

Qy       1176 ATCGGCCTCGGCGCCGGGCTCACCGTC----------------------------GTCGGC 1208
                  |||   :::||||||   :::   |||                           :::|||
Db        362 IleAlaIleGlyAlaMetIleGlyValLysPheGlyValAlaGlyThrGlyAsnIleGly 381

Qy       1209 CAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCC 1268
                  :::      ||||||         |||   :::|||
Db        382 Lys---SerAspAla-----------------AsnLeuIleValAlaLeuIleCysIle 394

Qy       1269 TACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATC 1328
                  |||||||||   |||:::         ||||||:::   |||:::      |||||||||
Db        395 TyrValAlaGlyPheAlaTrpSerTrpGlyProLeuGlyTrpLeuValProSerGluIle 414
```

```
Qy       1329 TTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACC 1388
Db        415 SerProLeuGluIleArgSerAlaAlaGlnAlaIleAsnValSerValAsnMetPhePhe 434

Qy       1389 AGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGC 1448
Db        435 ThrPheLeuValAlaGlnLeuPheLeuThrMetLeuCysHisMetLysPhe---GlyLeu 453

Qy       1449 TTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCG 1508
Db        454 PhePhePhePheAlaPhePheValValIleMetThrIlePheIleTyrLeuMetLeuPro 473

Qy       1509 GAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTC 1550
Db        474 GluThrLysAsnValProIleGluGluMetAsnArgValTrp 487

RESULT 15
US-09-679-686B-2
; Sequence 2, Application US/09679686B
; Patent No. 6624343
; GENERAL INFORMATION:
;  APPLICANT: Allen, Stephen M.
;  APPLICANT:  Lightner, Jonathan E.
;  APPLICANT:  Rafalski, J. Antoni
;  APPLICANT:  Thorpe, Catherine J.
;  TITLE OF INVENTION: HEXOSE CARRIER PROTEINS
;  FILE REFERENCE: BB1160 US NA
;  CURRENT APPLICATION NUMBER: US/09/679,686B
;  CURRENT FILING DATE:  2003-01-16
;  PRIOR APPLICATION NUMBER: 60/081,131
;  PRIOR FILING DATE: 1998-04-09
;  PRIOR APPLICATION NUMBER: PCT/US99/07561
;  PRIOR FILING DATE: 1999-04-07
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
;   LENGTH: 502
;   TYPE: PRT
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: (488)
;   OTHER INFORMATION: Xaa = any amino acid
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: (493)
;   OTHER INFORMATION: Xaa = any amino acid
US-09-679-686B-2

Alignment Scores:
Pred. No.:              8.01e-33        Length:         502
Score:                  533.00          Matches:        156
Percent Similarity:     48.34%          Conservative:   92
Best Local Similarity:  30.41%          Mismatches:     206
Query Match:            14.75%          Indels:         59
DB:                     4               Gaps:           15

US-10-051-902A-21 (1-2017) x US-09-679-686B-2 (1-502)

Qy         92 GGAGATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCG 151
Db          4 GlyPheGlyGlyGlyGlyAlaAlaAlaGlyArg-AlaGluLeuTyrGluGlyLysIleTh 23

Qy        152 G------TTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTA 205
Db         23 rGlyTyrPheIleLeuAlaCys---IleValGlySerPheGlyGlySerLeuPheGlyTy 42

Qy        206 CGATATCGGGGTGATGAGCGGGGCGTCG------------------------------ 233
Db         42 rAspLeuGlyValSerSerGlyValThrSerMetAspAspPheLeuValLysPhePhePr 62

Qy        234 ----CTGTACATCAAGAAG--------------------GACTTCAACATCAGTGACGG 268
Db         62 oAspValTyrArgArgLysGlnAlaHisLeuHisGluThrAspTyrCysLysTyrAspAs 82

Qy        269 GAAGGTGGAGGTTCTCATG---GGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGC 325
Db         82 nGlnValLeuThrLeuPheThrSerSerLeuTyrPheAlaGlyLeuValSerThrPheGl 102

Qy        326 GGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCAT 385
Db        102 yAlaSerTyrValThrLysArgHisGlyArgArgAlaSerIleMetGlyGlyAlaAlaSe 122

Qy        386 ATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGG 445
Db        122 rPhePheLeuGlyGlyAlaIleAsnGlyAlaAlaMetAsnIleAlaMetLeuIleValGl 142

Qy        446 CCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGC 505
Db        142 yArgIleLeuLeuGlyValGlyValGlyPheAlaAsnGlnAlaValProValTyrLeuSe 162

Qy        506 CGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAA 565
Db        162 rGluMetAlaProAlaArgLeuArgGlyMetLeuAsnIleGlyPheGlnLeuMetIleTh 182

Qy        566 CTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCT 625
Db        182 rIleGlyIleLeuAlaAlaGluLeuIleAsnTyrGlyThrAsnLysIleLysAlaGlyTy 202

Qy        626 CGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGT 685
Db        202 rGlyTrpArgValSerLeuAlaLeuAlaAlaValProAlaAlaIleIleThrLeuGlySe 222

Qy        686 GCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAA 745
Db        222 rLeuPheLeuProAspThrProAsnSerLeuLeuGluArgGlyHisProGluGluAlaAr 242

Qy        746 GGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACAT 805
Db        242 gArgMetLeuArgArgIleArgGlyThr---AspAspIleGlyGluGluTyrAlaAspLe 261

Qy        806 CAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAG 865
Db        261 uValAlaAlaSer---------GluGlu------------------------------ 267

Qy        866 AGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCAT 925
Db        268 ----AlaArgGlnValArgHisProTrpArgAsnIleLeu---------ArgArgArgTy 283

Qy        926 GCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCA 985
Db        283 rArgAlaGlnLeuThrMetAlaValAlaIleProPhePheGlnGlnLeuLysGlyIleAs 303

Qy        986 CTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACA 1045
Db        303 nValIleMetPheTyrAlaProValLeuPheAspThrLeuGlyPheLysLysGluAla-- 322

Qy       1046 CTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGAC 1105
Db        323 -PheLeuMetSerSerValIleThrGlyLeuValAsnValPheAlaThrValValSerIl 342

Qy       1106 TTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGG----- 1160
Db        342 eValThrValAspArgValGlyArgArgLysLeuPheLeu------GlnGlyGlyAlaGl 360

Qy       1161 ------------------ATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGT 1201
Db        360 nMetIleValCysGlnLeuIleValGlyThrLeuIleAlaAlaLysPheGlyThrSerGl 380

Qy       1202 CGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCAC 1261
Db        380 yThrGly---------AspIleAlaLysGlyTyrAla---AlaValValValValPheIl 396

Qy       1262 CCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTC 1321
Db        396 eCysAlaTyrValAlaGlyPheAlaTrpSerTrpGlyProLeuGlyTrpLeuValProSe 416
```

```
Qy       1322 GGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCG 1381
              ||||||||||||||||::::::|||      |||   |||:::   |||:::   |||
Db        416 rGluIlePheProLeuGluIleArgProAlaGlyGlnSerIleAsnValSerValAsnMe 436

Qy       1382 CGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGG 1441
                  :::     |||:::      ||||||::::::
Db        436 tPhePheThrPheCysIleAlaGlnAlaPheLeuThrMetLeuCysHisPheLysPhe-- 455

Qy       1442 CGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTA 1501
               |||   |||:::   ::::::|||         :::      ||||||         ::
Db        456 -GlyLeuPheTyrPhePheAlaGlyTrpValValIleMetThrValPheIleAlaPhePh 475

Qy       1502 CCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATG 1538
              :||||||||||||:::         :::||||||||
Db        475 eLeuProGluThrLysAsnValProIleGluGluMet 487

Search completed: June 30, 2004, 18:54:21
Job time : 64 secs
```

GenCore version 5.1.6

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 30, 2004, 18:42:00 ; Search time 38 Seconds
(without alignments)
10211.490 Million cell updates/sec

Title: US-10-051-902A-21
Perfect score: 3614
Sequence: 1 cttacatgtaagctcgtgcc.....aaaaaaaaaaaaaaaaaaa 2017

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 566732

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10051902/runat_30062004_164723_20424/app_query.fasta_1.2183
-DB=PIR_78 -QFMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10051902_@CGN_1_1_52_@runat_30062004_164723_20424 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1589	44.0	549	2	T14606	probable sugar tra
2	1579	43.7	511	2	H84536	probable sugar tra
3	1559	43.1	511	2	A84537	probable sugar tra
4	1385.5	38.3	508	2	G84564	probable sugar tra
5	1361	37.7	493	2	A85433	sugar transporter
6	1143	31.6	547	2	C84593	probable sugar tra
7	673.5	18.6	521	2	G84864	probable membrane
8	670.5	18.6	457	2	E70070	metabolite transpo
9	629.5	17.4	582	2	F71431	hypothetical prote
10	626	17.3	580	2	D86426	hypothetical prote
11	609.5	16.9	461	2	D70073	metabolite transpo
12	609.5	16.9	473	2	G69789	sugar transporter
13	601	16.6	557	2	T38125	myo-inositol trans
14	586.5	16.2	580	2	D84772	probable sugar tra
15	584.5	16.2	472	2	S47089	arabinose-proton s
16	584.5	16.2	560	2	T51485	sugar transporter-
17	581	16.1	606	2	T27072	hypothetical prote
18	579.5	16.0	472	2	B26430	L-arabinose isomer
19	579.5	16.0	472	2	B91091	L-arabinose isomer
20	579.5	16.0	472	2	E85936	L-arabinose isomer
21	578.5	16.0	464	2	C91106	galactose-proton s
22	578.5	16.0	464	2	F85951	galactose-proton s
23	576	15.9	487	2	E96782	hypothetical prote
24	576	15.9	584	2	S69555	myo-inositol trans
25	574	15.9	491	2	A26430	xylose transport p
26	574	15.9	491	2	F91255	xylose-proton symp
27	574	15.9	491	2	B86096	xylose-proton symp
28	572.5	15.8	464	2	F65079	galactose-proton s
29	572.5	15.8	612	2	B40538	myo-inositol trans
30	572	15.8	519	2	C87350	major facilitator
31	571	15.8	464	2	AC0877	galactose-proton s
32	568.5	15.7	471	2	AB0868	L-arabinose isomer
33	564	15.6	502	2	B70845	probable sygar tra
34	562	15.6	639	2	T23658	hypothetical prote
35	560	15.5	517	2	D96539	hypothetical prote
36	555.5	15.4	482	2	B69803	metabolite transpo
37	552	15.3	613	2	T27077	hypothetical prote
38	543.5	15.0	522	2	S12042	glucose transport
39	543.5	15.0	522	2	E86246	glucose transporte
40	542	15.0	490	2	T14545	probable sugar tra
41	539	14.9	433	2	G86812	D-xylose proton-sy
42	539	14.9	534	2	S38435	hexose transport p
43	538.5	14.9	534	2	S14144	hexose transport p
44	536	14.8	514	2	S25009	monosaccharide tra
45	534	14.8	472	2	T35662	probable sugar tra

ALIGNMENTS

RESULT 1
T14606
probable sugar transport protein 205 - beet
C;Species: Beta vulgaris (beet)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 01-Dec-2000
C;Accession: T14606; T14617
R;Chiou, T.J.; Bush, D.R.
Plant Physiol. 110, 511-520, 1996
A;Title: Molecular cloning, immunochemical localization to the vacuole, and expression ir
A;Reference number: Z18131; MUID:96351183; PMID:8742332
A;Accession: T14606
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-549 <CHI>
A;Cross-references: EMBL:U64902; NID:g1778092; PID:g1778093
A;Accession: T14617
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-541,'SVQV' <CH2>
A;Cross-references: EMBL:U64903; NID:g1778094; PID:g1778095
C;Genetics:
A;Note: BvcDNA-205; BvcDNA-397
C;Superfamily: glucose transport protein
C;Keywords: sugar transport

Alignment Scores:

Pred. No.:	3.74e-85	Length:	549
Score:	1589.00	Matches:	326
Percent Similarity:	72.23%	Conservative:	59
Best Local Similarity:	61.16%	Mismatches:	126
Query Match:	43.97%	Indels:	22
DB:	2	Gaps:	5

US-10-051-902A-21 (1-2017) x T14606 (1-549)

```
Qy          78 TCCTCTGCACCACCGGAGATGGCTTCCGCCGCGCTG----------CCGGAGGCCGTC 125
               |||      ||||||      ||||||      :::           |||
Db          10 SerAspProProProThrThrAlaSerLysValIleAlaAspPheAspProLeuLysLys 29
```

```
Qy      126 GCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATG 185
Db       30 ProProLysArgAsn---------LysPheAlaPheAlaCysAlaThrLeuAlaSerMet 46

Qy      186 ACCTCCATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAG 245
Db       47 ThrSerValLeuLeuGlyTyrAspIleGlyValMetSerGlyAlaIleIleTyrLeuLys 66

Qy      246 AAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTAC 305
Db       67 GluAspTrpHisIleSerAspThrGlnIleGlyValLeuValGlyIleLeuAsnIleTyr 86

Qy      306 TCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACC 365
Db       87 CysLeuPheGlySerPheAlaAlaGlyArgThrSerAspTrpIleGlyArgArgTyrThr 106

Qy      366 ATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAAC 425
Db      107 IleValLeuAlaGlyAlaIlePhePheValGlyAlaLeuLeuMetGlyPheAlaThrAsn 126

Qy      426 TACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATG 485
Db      127 TyrAlaPheLeuMetValGlyArgPheValThrGlyIleGlyValGlyTyrAlaLeuMet 146

Qy      486 ATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCG 545
Db      147 IleAlaProValTyrThrAlaGluValSerProAlaSerSerArgGlyPheLeuThrSer 166

Qy      546 TTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTC 605
Db      167 PheProGluValPheIleAsnAlaGlyIleLeuLeuGlyTyrIleSerAsnLeuAlaPhe 186

Qy      606 TCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCC 665
Db      187 SerSerLeuProThrHisLeuSerTrpArgPheMetLeuGlyIleGlyAlaIleProSer 206

Qy      666 GTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAG 725
Db      207 IlePheLeuAlaIleGlyValLeuAlaMetProGluSerProArgTrpLeuValMetGln 226

Qy      726 GGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCC 785
Db      227 GlyArgLeuGlyAspAlaLysLysValLeuAsnArgIleSerAspSerProGluGluAla 246

Qy      786 GCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGAC 845
Db      247 GlnLeuArgLeuSerGluIleLysGlnThrAlaGlyIleProAlaGluCysAspGluAsp 266

Qy      846 GTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATC 905
Db      267 IleTyrLysValGluLysThrLysIleLysSerGlyAsnAlaValTrpLysGluLeuPhe 286

Qy      906 CTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGGATCGGCATCCACTTCTTC 965
Db      287 PheAsnProThrProAlaValArgArgAlaValIleAlaGlyIleGlyIleHisPhePhe 306

Qy      966 CAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCC 1025
Db      307 GlnGlnAlaSerGlyIleAspAlaValValLeuTyrSerProArgIlePheGlnSerAla 326

Qy     1026 GGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGG 1085
Db      327 GlyIleThrAsnAlaArgLysGlnLeuLeuAlaThrValAlaValGlyValValLysThr 346

Qy     1086 CTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTG 1145
Db      347 LeuPheIleLeuValAlaThrPheGlnLeuAspLysTyrGlyArgArgProLeuLeuLeu 366

Qy     1146 GGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTC 1205
Db      367 ThrSerValGlyGlyMetIleIleAlaIleLeuThrLeuAlaMetSerLeuThrValIle 386

Qy     1206 GGCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTC 1265
Db      387 ---AspHisSerHisHisLysIleThrTrpAlaIleAlaLeuCysIleThrMetValCys 405

Qy     1266 GCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAG 1325
Db      406 AlaValValAlaSerPheSerIleGlyLeuGlyProIleThrTrpValTyrSerSerGlu 425

Qy     1326 ATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTC 1385
Db      426 ValPheProLeuArgLeuArgAlaGlnGlyThrSerMetGlyValAlaValAsnArgVal 445

Qy     1386 ACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGC 1445
Db      446 ValSerGlyValIleSerIlePhePheLeuProLeuSerHisLysIleThrThrGlyGly 465

Qy     1446 AGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTC 1505
Db      466 AlaPhePheLeuPheGlyGlyIleAlaIleIleAlaTrpPhePhePheLeuThrPheLeu 485

Qy     1506 CCGGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGAC--------- 1556
Db      486 ProGluThrArgGlyArgThrLeuGluAsnMetHisGluLeuPheGluAspPheArgTrp 505

Qy     1557 ------------ACGGCCGCCGCCTCGGAATCAGACGAGCCAGCC-------------- 1589
Db      506 ArgGluSerPheProGlyAsnLysSerAsnAsnAspGluAsnSerThrArgLysGlnSer 525

Qy     1590 ------AAGGAGAAGAAGAAGGTGGAAATGGCCGCCACT 1622
Db      526 AsnGlyAsnAspLysSerGlnValGlnLeuGlyGluThr 538
```

RESULT 2
H84536
probable sugar transporter [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C;Accession: H84536
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: H84536
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-511 <STO>
A;Cross-references: GB:AE002093; NID:g4678208; PIDN:AAD26954.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g16120
A;Map position: 2
C;Superfamily: glucose transport protein

Alignment Scores:

Pred. No.:	1.42e-84	Length:	511
Score:	1579.00	Matches:	306
Percent Similarity:	75.10%	Conservative:	77
Best Local Similarity:	60.00%	Mismatches:	121
Query Match:	43.69%	Indels:	6
DB:	2	Gaps:	2

US-10-051-902A-21 (1-2017) x H84536 (1-511)

```
Qy        96 ATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCG--------------CCGAAGAAGAAG 140
Db         1 MetAsnSerSerGlyValGluGlnGlyValValIleAlaGluSerGluProProArgGly 20

Qy       141 GGCAACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTC 200
Db        21 AsnArgSerArgTyrAlaPheAlaCysAlaIleLeuAlaSerMetThrSerIleIleLeu 40

Qy       201 GGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATC 260
```

```
Db        41 GlyTyrAspIleGlyValMetSerGlyAlaSerIlePheIleLysAspAspLeuLysLeu 60

Qy       261 AGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCC 320
Db        61 SerAspValGlnLeuGluIleLeuMetGlyIleLeuAsnIleTyrSerLeuValGlySer 80

Qy       321 TTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCC 380
Db        81 GlyAlaAlaGlyArgThrSerAspTrpLeuGlyArgArgTyrThrIleValLeuAlaGly 100

Qy       381 GTCATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATG 440
Db       101 AlaPhePhePheCysGlyAlaLeuLeuMetGlyPheAlaThrAsnTyrProPheIleMet 120

Qy       441 TTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTAC 500
Db       121 ValGlyArgPheValAlaGlyIleGlyValGlyTyrAlaMetMetIleAlaProValTyr 140

Qy       501 ACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTC 560
Db       141 ThrAlaGluValAlaProAlaSerSerArgGlyPheLeuThrSerPheProGluIlePhe 160

Qy       561 ATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTG 620
Db       161 IleAsnIleGlyIleLeuLeuGlyTyrValSerAsnTyrPhePheSerLysLeuProGlu 180

Qy       621 AACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTC 680
Db       181 HisLeuGlyTrpArgPheMetLeuGlyValGlyAlaValProSerValPheLeuAlaIle 200

Qy       681 ATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGAC 740
Db       201 GlyValLeuAlaMetProGluSerProArgTrpLeuValLeuGlnGlyArgLeuGlyAsp 220

Qy       741 GCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCC 800
Db       221 AlaPheLysValLeuAspLysThrSerAsnThrLysGluGluAlaIleSerArgLeuAsp 240

Qy       801 GACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCC 860
Db       241 AspIleLysArgAlaValGlyIleProAspAspMetThrAspAspValIleValValPro 260

Qy       861 AAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCG 920
Db       261 AsnLysLysSerAlaGly---LysGlyValTrpLysAspLeuLeuValArgProThrPro 279

Qy       921 GCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGC 980
Db       280 SerValArgHisIleLeuIleAlaCysLeuGlyIleHisPheAlaGlnGlnAlaSerGly 299

Qy       981 ATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGAC 1040
Db       300 IleAspAlaValValLeuTyrSerProThrIlePheSerLysAlaGlyLeuLysSerLys 319

Qy      1041 AAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTG 1100
Db       320 AsnAspGlnLeuLeuAlaThrValAlaValGlyValValLysThrLeuPheIleValVal 339

Qy      1101 GCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGG 1160
Db       340 GlyThrCysValValAspArgPheGlyArgArgAlaLeuLeuLeuThrSerMetGlyGly 359

Qy      1161 ATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGAC 1220
Db       360 MetPheLeuSerLeuThrAlaLeuGlyThrSerLeuThrValIleAsnArgAsnProGly 379

Qy      1221 GCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTC 1280
Db       380 GlnThrLeuLysTrpAlaIleGlyLeuAlaValThrThrValMetThrPheValAlaThr 399

Qy      1281 TTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAG 1340

Db       400 PheSerIleGlyAlaGlyProValThrTrpValTyrCysSerGluIlePheProValArg 419

Qy      1341 GTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATC 1400
Db       420 LeuArgAlaGlnGlyAlaSerLeuGlyValMetLeuAsnArgLeuMetSerGlyIleIle 439

Qy      1401 TCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTAC 1460
Db       440 GlyMetThrPheLeuSerLeuSerLysGlyLeuThrIleGlyGlyAlaPheLeuLeuPhe 459

Qy      1461 TCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGC 1520
Db       460 AlaGlyValAlaAlaAlaAlaTrpValPhePhePheThrPheLeuProGluThrArgGly 479

Qy      1521 CGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGAC 1580
Db       480 IleProLeuGluGluMetGluThrLeuPheGlySerTyrThrAlaAsnLysLysAsnAsn 499

Qy      1581 GAGCCAGCCAAGGAGAAGAAGAAGGTGGAA 1610
Db       500 SerMetSerLysAspAsnGluValValAsp 509
```

RESULT 3
A84537
probable sugar transporter [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C;Accession: A84537
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: A84537
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-511 <STO>
A;Cross-references: GB:AE002093; NID:g4678209; PIDN:AAD26955.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g16130
A;Map position: 2
C;Superfamily: glucose transport protein

Alignment Scores:

Pred. No.:	2.08e-83	Length:	511
Score:	1559.00	Matches:	305
Percent Similarity:	75.15%	Conservative:	79
Best Local Similarity:	59.69%	Mismatches:	123
Query Match:	43.14%	Indels:	4
DB:	2	Gaps:	3

US-10-051-902A-21 (1-2017) x A84537 (1-511)

```
Qy        81 TCTGCACCACCGGAGATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAG 140
Db         2 SerSerSerGlyGluGluArgGlyValValValAlaGluSerGluProPro------Arg 19

Qy       141 GGCAAC---GTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTC 197
Db        20 GlyAsnArgSerArgPheAlaPheAlaCysAlaIleLeuAlaSerMetThrSerIleIle 39

Qy       198 CTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAAC 257
Db        40 LeuGlyTyrAspIleGlyValMetSerGlyAlaAlaIlePheIleLysAspAspLeuLys 59

Qy       258 ATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGC 317
Db        60 LeuSerAspValGlnLeuGluIleLeuMetGlyIleLeuAsnIleTyrSerLeuIleGly 79

Qy       318 TCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCC 377
Db        80 SerGlyAlaAlaGlyArgThrSerAspTrpIleGlyArgArgTyrThrIleValLeuAla 99
```

```
Qy        378 GCCGTCATATTCTTCGCGGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTC 437
Db        100 GlyPhePhePhePheCysGlyAlaLeuLeuMetGlyPheAlaThrAsnTyrProPheIle 119

Qy        438 ATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTG 497
Db        120 MetValGlyArgPheValAlaGlyIleGlyValGlyTyrAlaMetMetIleAlaProVal 139

Qy        498 TACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTG 557
Db        140 TyrThrThrGluValAlaProAlaSerSerArgGlyPheLeuSerSerPheProGluIle 159

Qy        558 TTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCG 617
Db        160 PheIleAsnIleGlyIleLeuLeuGlyTyrValSerAsnTyrPhePheAlaLysLeuPro 179

Qy        618 CTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCG 677
Db        180 GluHisIleGlyTrpArgPheMetLeuGlyIleGlyAlaValProSerValPheLeuAla 199

Qy        678 CTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCG 737
Db        200 IleGlyValLeuAlaMetProGluSerProArgTrpLeuValMetGlnGlyArgLeuGly 219

Qy        738 GACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTG 797
Db        220 AspAlaPheLysValLeuAspLysThrSerAsnThrLysGluGluAlaIleSerArgLeu 239

Qy        798 GCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTC 857
Db        240 AsnAspIleLysArgAlaValGlyIleProAspAspMetThrAspAspValIleValVal 259

Qy        858 CCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACC 917
Db        260 ProAsnLysLysSerAlaGly---LysGlyValTrpLysAspLeuLeuValArgProThr 278

Qy        918 CCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTG 977
Db        279 ProSerValArgHisIleLeuIleAlaCysLeuGlyIleHisPheSerGlnGlnAlaSer 298

Qy        978 GGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAAC 1037
Db        299 GlyIleAspAlaValValLeuTyrSerProThrIlePheSerArgAlaGlyLeuLysSer 318

Qy       1038 GACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTG 1097
Db        319 LysAsnAspGlnLeuLeuAlaThrValAlaValGlyValValLysThrLeuPheIleVal 338

Qy       1098 TTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGC 1157
Db        339 ValGlyThrCysLeuValAspArgPheGlyArgArgAlaLeuLeuLeuThrSerMetGly 358

Qy       1158 GGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCC 1217
Db        359 GlyMetPhePheSerLeuThrAlaLeuGlyThrSerLeuThrValIleAspArgAsnPro 378

Qy       1218 GACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCC 1277
Db        379 GlyGlnThrLeuLysTrpAlaIleGlyLeuAlaValThrThrValMetThrPheValAla 398

Qy       1278 TTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTC 1337
Db        399 ThrPheSerLeuGlyAlaGlyProValThrTrpValTyrAlaSerGluIlePheProVal 418

Qy       1338 CAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTC 1397
Db        419 ArgLeuArgAlaGlnGlyAlaSerLeuGlyValMetLeuAsnArgLeuMetSerGlyIle 438

Qy       1398 ATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTC 1457
Db        439 IleGlyMetThrPheLeuSerLeuSerLysGlyLeuThrIleGlyGlyAlaPheLeuLeu 458

Qy       1458 TACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGC 1517
Db        459 PheAlaGlyValAlaValAlaAlaTrpValPhePhePheThrPheLeuProGluThrArg 478

Qy       1518 GGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCA 1577
Db        479 GlyValProLeuGluGluIleGluSerLeuPheGlySerTyrSerAlaAsnLysLysAsn 498

Qy       1578 GACGAGCCAGCCAAGGAGAAGAAGAAGGTGGAA 1610
Db        499 AsnValMetSerLysGlyLysGlnValValAsp 509
```

RESULT 4
G84564
probable sugar transporter [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C;Accession: G84564
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: G84564
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-508 <STO>
A;Cross-references: GB:AE002093; NID:g4218010; PIDN:AAD12218.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g18480
A;Map position: 2
C;Superfamily: glucose transport protein

Alignment Scores:

Pred. No.:	2.7e-73	Length:	508
Score:	1385.50	Matches:	271
Percent Similarity:	72.41%	Conservative:	78
Best Local Similarity:	56.22%	Mismatches:	120
Query Match:	38.34%	Indels:	13
DB:	2	Gaps:	3

US-10-051-902A-21 (1-2017) x G84564 (1-508)

```
Qy        150 CGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGAT 209
Db         20 LysPheAlaPheGlyCysAlaIleValAlaSerIleIleSerIleIlePheGlyTyrAsp 39

Qy        210 ATCGGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGG 269
Db         40 ThrGlyValMetSerGlyAlaGlnIlePheIleArgAspAspLeuLysIleAsnAspThr 59

Qy        270 AAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCG 329
Db         60 GlnIleGluValLeuAlaGlyIleLeuAsnLeuCysAlaLeuValGlySerLeuThrAla 79

Qy        330 GGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTC 389
Db         80 GlyLysThrSerAspValIleGlyArgArgTyrThrIleAlaLeuSerAlaValIlePhe 99

Qy        390 TTCGCGGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGC 449
Db        100 LeuValGlySerValLeuMetGlyTyrGlyProAsnTyrProValLeuMetValGlyArg 119

Qy        450 TTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAG 509
Db        120 CysIleAlaGlyValGlyValGlyPheAlaLeuMetIleAlaProValTyrSerAlaGlu 139

Qy        510 GTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTC 569
Db        140 IleSerSerAlaSerHisArgGlyPheLeuThrSerLeuProGluLeuCysIleSerLeu 159

Qy        570 GGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGG 629
```

```
Db        160 GlyIleLeuLeuGlyTyrValSerAsnTyrCysPheGlyLysLeuThrLeuLysLeuGly 179

Qy        630 TGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTC 689

Db        180 TrpArgLeuMetLeuGlyIleAlaAlaProSerLeuIleLeuAlaPheGlyIleThr 199

Qy        690 GGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTG 749

Db        200 ArgMetProGluSerProArgTrpLeuValMetGlnGlyArgLeuGluGluAlaLysLys 219

Qy        750 GTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAG 809

Db        220 IleMetValLeuValSerAsnThrGluGluGluAlaGluGluArgPheArgAspIleLeu 239

Qy        810 GCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGG 869

Db        240 ThrAlaAlaGluVal----------------AspValThrGluIleLysGluValGly 253

Qy        870 AGCGGAAACGAGAAGCGG--------------GTGTGGAAGGAGCTCATCCTGTCCCCG 914

Db        254 GlyGlyValLysLysLysAsnHisGlyLysSerValTrpArgGluLeuValIleLysPro 273

Qy        915 ACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCG 974

Db        274 ArgProAlaValArgLeuIleLeuIleAlaAlaValGlyIleHisPhePheGluHisAla 293

Qy        975 TTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACG 1034

Db        294 ThrGlyIleGluAlaValValLeuTyrSerProArgIlePheLysLysAlaGlyValVal 313

Qy       1035 AACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATC 1094

Db        314 SerLysAspLysLeuLeuLeuAlaThrValGlyValGlyLeuThrLysAlaPhePheIle 333

Qy       1095 TTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACG 1154

Db        334 IleIleAlaThrPheLeuLeuAspLysValGlyArgArgLysLeuLeuLeuThrSerThr 353

Qy       1155 GGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCAC 1214

Db        354 GlyGlyMetValPheAlaLeuThrSerLeuAlaValSerLeuThrMetValGlnArgPhe 373

Qy       1215 CCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTC 1274

Db        374 ------GlyArgLeuAlaTrpAlaLeuSerLeuSerIleValSerThrTyrAlaPheVal 391

Qy       1275 GCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCG 1334

Db        392 AlaPhePheSerIleGlyLeuGlyProIleThrTrpValTyrSerSerGluIlePhePro 411

Qy       1335 CTCCAGGTGCGCGGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGC 1394

Db        412 LeuArgLeuArgAlaGlnGlyAlaSerIleGlyValAlaValAsnArgIleMetAsnAla 431

Qy       1395 GTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTC 1454

Db        432 ThrValSerMetSerPheLeuSerMetThrLysAlaIleThrThrGlyGlyValPhePhe 451

Qy       1455 CTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACC 1514

Db        452 ValPheAlaGlyIleAlaValAlaAlaTrpTrpPhePhePhePheMetLeuProGluThr 471

Qy       1515 CGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAA 1574

Db        472 LysGlyLeuProLeuGluGluMetGluLysLeuPheGlyGlyGlyGlyProArgGlyAsp 491

Qy       1575 TCAGAC 1580

Db        492 ArgAsp 493

RESULT 5
A85433
```

```
sugar transporter like protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 02-Mar-2001
C;Accession: A85433
R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sprin
Nature 402, 769-777, 1999
A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A;Reference number: A85001; MUID:20083488; PMID:10617198
A;Accession: A85433
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-493 <STO>
A;Cross-references: GB:NC_001268; NID:g7270615; PIDN:CAB80333.1; GSPDB:GN00140
C;Genetics:
A;Gene: AT4g36670
A;Map position: 4
C;Superfamily: glucose transport protein

Alignment Scores:
Pred. No.:                7.21e-72        Length:          493
Score:                    1361.00         Matches:         271
Percent Similarity:       72.95%          Conservative:    85
Best Local Similarity:    55.53%          Mismatches:      120
Query Match:              37.66%          Indels:          12
DB:                       2               Gaps:            5

US-10-051-902A-21 (1-2017) x A85433 (1-493)

Qy        132 AAGAAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCC 191

Db          9 LysProAlaGlyValAsnArgPheAlaLeuGlnCysAlaIleValAlaSerIleValSer 28

Qy        192 ATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGAC 251

Db         29 IleIlePheGlyTyrAspThrGlyValMetSerGlyAlaMetValPheIleGluGluAsp 48

Qy        252 TTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTC 311

Db         49 LeuLysThrAsnAspValGlnIleGluValLeuThrGlyIleLeuAsnLeuCysAlaLeu 68

Qy        312 ATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTG 371

Db         69 ValGlySerLeuLeuAlaGlyArgThrSerAspIleIleGlyArgArgTyrThrIleVal 88

Qy        372 TTCGCCGCCGTCATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCC 431

Db         89 LeuAlaSerIleLeuPheMetLeuGlySerIleLeuMetGlyTrpGlyProAsnTyrPro 108

Qy        432 ATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCG 491

Db        109 ValLeuLeuSerGlyArgCysThrAlaGlyLeuGlyValGlyPheAlaLeuMetValAla 128

Qy        492 CCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCG 551

Db        129 ProValTyrSerAlaGluIleAlaThrAlaSerHisArgGlyLeuLeuAlaSerLeuPro 148

Qy        552 GAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGC 611

Db        149 HisLeuCysIleSerIleGlyIleLeuLeuGlyTyrIleValAsnTyrPhePheSerLys 168

Qy        612 TTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTG 671

Db        169 LeuProMetHisIleGlyTrpArgLeuMetLeuGlyIleAlaAlaValProSerLeuVal 188

Qy        672 CTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGC 731

Db        189 LeuAlaPheGlyIleLeuLysMetProGluSerProArgTrpLeuIleMetGlnGlyArg 208

Qy        732 CTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAG 791

Db        209 LeuLysGluGlyLysGluIleLeuGluLeuValSerAsnSerProGluGluAlaGluLeu 228

Qy        792 CGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTG 851
```

```
Db        229 ArgPheGlnAspIleLysAlaAlaAlaGlyIleAspProLysCysValAspAspValVal 248

Qy        852 ACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCC 911
Db        249 LysMetGluGlyLysLysThrHisGlyGlu---GlyValTrpLysGluLeuIleLeuArg 267

Qy        912 CCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCAT 971
Db        268 ProThrProAlaValArgArgValLeuLeuThrAlaLeuGlyIleHisPhePheGlnHis 287

Qy        972 GCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTA 1031
Db        288 AlaSerGlyIleGluAlaValLeuLeuTyrGlyProArgIlePheLysLysAlaGlyIle 307

Qy       1032 ---ACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTT 1088
Db        308 ThrThrLysAspLysLeuPheLeu---ValThrIleGlyValGlyIleMetLysThrThr 326

Qy       1089 TTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGC 1148
Db        327 PheIlePheThrAlaThrLeuLeuLeuAspLysValGlyArgArgLysLeuLeuLeuThr 346

Qy       1149 AGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGC 1208
Db        347 SerValGlyGlyMetValIleAlaLeuThrMetLeuGlyPheGlyLeuThr---MetAla 365

Qy       1209 CAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCC 1268
Db        366 GlnAsnAlaGlyGlyLysLeuAlaTrpAlaLeuValLeuSerIleValAlaAlaTyrSer 385

Qy       1269 TACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATC 1328
Db        386 PheValAlaPhePheSerIleGlyLeuGlyProIleThrTrpValTyrSerSerGluVal 405

Qy       1329 TTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACC 1388
Db        406 PheProLeuLysLeuArgAlaGlnGlyAlaSerLeuGlyValAlaValAsnArgValMet 425

Qy       1389 AGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGC 1448
Db        426 AsnAlaThrValSerMetSerPheLeuSerLeuThrSerAlaIleThrThrGlyGlyAla 445

Qy       1449 TTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCG 1508
Db        446 PhePheMetPheAlaGlyValAlaAlaValAlaTrpAsnPhePhePhePheLeuLeuPro 465

Qy       1509 GAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTC------------------ 1550
Db        466 GluThrLysGlyLysSerLeuGluGluIleGluAlaLeuPheGlnArgAspGlyAspLys 485

Qy       1551 ------GGCGACACGGCCGCCGCC 1568
Db        486 ValArgGlyGluAsnGlyAlaAla 493
```

RESULT 6
C84593
probable sugar transporter [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 03-Jun-2002
C;Accession: C84593
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: C84593
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-547 <STO>
A;Cross-references: GB:AE002093; NID:g4454470; PIDN:AAD20917.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g20780
A;Map position: 2
C;Superfamily: glucose transport protein

Alignment Scores:

Pred. No.:	3.71e-59	Length:	547
Score:	1143.00	Matches:	234
Percent Similarity:	65.85%	Conservative:	88
Best Local Similarity:	47.85%	Mismatches:	125
Query Match:	31.63%	Indels:	42
DB:	2	Gaps:	5

US-10-051-902A-21 (1-2017) x C84593 (1-547)

```
Qy        150 CGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTAC--- 206
Db         53 LysTyrValMetAlaCysAlaPhePheAlaSerLeuAsnAsnValLeuLeuGlyTyrGly 72

Qy        206 ------------------------------------------------------------ 206

Db         73 ArgPheTyrLeuTyrAsnArgIleLeuLeuLeuLeuLeuTyrPheValAspLeuGlnLys 92

Qy        207 GATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGAC 266
Db         93 AspValGlyValMetSerGlyAlaValLeuPheIleGlnGlnAspLeuLysIleThrGlu 112

Qy        267 GGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCG 326
Db        113 ValGlnThrGluValLeuIleGlySerLeuSerIleIleSerLeuPheGlySerLeuAla 132

Qy        327 GCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATA 386
Db        133 GlyGlyArgThrSerAspSerIleGlyArgLysTrpThrMetAlaLeuAlaAlaLeuVal 152

Qy        387 TTCTTCGCGGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGC 446
Db        153 PheGlnThrGlyAlaAlaValMetAlaValAlaProSerPheGluValLeuMetIleGly 172

Qy        447 CGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCC 506
Db        173 ArgThrLeuAlaGlyIleGlyIleGlyLeuGlyValMetIleAlaProValTyrIleAla 192

Qy        507 GAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAAC 566
Db        193 GluIleSerProThrValAlaArgGlyPhePheThrSerPheProGluIlePheIleAsn 212

Qy        567 TTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTC 626
Db        213 LeuGlyIleLeuLeuGlyTyrValSerAsnTyrAlaPheSerGlyLeuSerValHisIle 232

Qy        627 GGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTG 686
Db        233 SerTrpArgIleMetLeuAlaValGlyIleLeuProSerValPheIleGlyPheAlaLeu 252

Qy        687 CTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAG 746
Db        253 CysValIleProGluSerProArgTrpLeuValMetLysGlyArgValAspSerAlaArg 272

Qy        747 GTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATC 806
Db        273 GluValLeuMetLysThrAsnGluArgAspAspGluAlaGluGluArgLeuAlaGluIle 292

Qy        807 AAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGA 866
Db        293 GlnLeuAlaAlaAlaHisThrGlu------------------------------------ 300

Qy        867 GGGAGCGGAAACGAGAAGCGG---GTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCC 923
Db        301 ------GlySerGluAspArgProValTrpArgGluLeu---LeuSerProSerProVal 317

Qy        924 ATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATT 983
Db        318 ValArgLysMetLeuIleValGlyPheGlyIleGlnCysPheGlnGlnIleThrGlyIle 337
```

```
Qy        984 CACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAA 1043
Db        338 AspAlaThrValTyrTyrSerProGluIleLeuLysGluAlaGlyIleGlnAspGluThr 357

Qy       1044 CACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCG 1103
Db        358 LysLeuLeuAlaAlaThrValAlaValGlyValThrLysThrValPheIleLeuPheAla 377

Qy       1104 ACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATA 1163
Db        378 ThrPheLeuIleAspSerValGlyArgLysProLeuLeuTyrValSerThrIleGlyMet 397

Qy       1164 ATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCC 1223
Db        398 ThrLeuCysLeuPheCysLeuSerPheThrLeuThrPheLeuGlyGlnGly--------- 414

Qy       1224 AAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTC 1283
Db        415 ------ThrLeuGlyIleThrLeuAlaLeuLeuPheValCysGlyAsnValAlaPhePhe 432

Qy       1284 TCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTG 1343
Db        433 SerIleGlyMetGlyProValCysTrpValLeuThrSerGluIlePheProLeuArgLeu 452

Qy       1344 CGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCC 1403
Db        453 ArgAlaGlnAlaSerAlaLeuGlyAlaValGlyAsnArgValCysSerGlyLeuValAla 472

Qy       1404 ATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCC 1463
Db        473 MetSerPheLeuSerValSerArgAlaIleThrValGlyGlyThrPhePheValPheSer 492

Qy       1464 GGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGG 1523
Db        493 LeuValSerAlaLeuSerValIlePheValTyrValLeuValProGluThrSerGlyLys 512

Qy       1524 ACGCTGGAGGAGATGAGCAAGCTGTTC 1550
Db        513 SerLeuGluGlnIleGluLeuMetPhe 521
```

RESULT 7
G84864
probable membrane transporter [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C;Accession: G84864
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L. euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: G84864
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-521 <STO>
A;Cross-references: GB:AE002093; NID:g2289003; PIDN:AAB64332.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g43330
A;Map position: 2
C;Superfamily: glucose transport protein

Alignment Scores:

Pred. No.:	8.59e-32	Length:	521
Score:	673.50	Matches:	160
Percent Similarity:	51.68%	Conservative:	101
Best Local Similarity:	31.68%	Mismatches:	177
Query Match:	18.64%	Indels:	67
DB:	2	Gaps:	8

US-10-051-902A-21 (1-2017) x G84864 (1-521)

```
Qy        141 GGCAACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTC 200
Db         27 GlyAsn---SerTyrIleLeuGlyLeuThrValThrAlaGlyIleGlyGlyLeuLeuPhe 45

Qy        201 GGCTACGATATCGGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATC 260
Db         46 GlyTyrAspThrGlyValIleSerGlyAlaLeuLeuTyrIleLysAspAspPheGluVal 65

Qy        261 ----------------------------------------------AGTGACGGGAAGGTG 275
Db         66 ValLysGlnSerSerPheLeuGlnValTyrAsnValSerSerPheThrSerSerLysLeu 85

Qy        276 GAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGG 335
Db         86 GluThrIleValSerMetAlaLeuValGlyAlaMetIleGlyAlaAlaAlaGlyGlyTrp 105

Qy        336 ACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCG 395
Db        106 IleAsnAspTyrTyrGlyArgLysLysAlaThrLeuPheAlaAspValValPheAlaAla 125

Qy        396 GGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTG 455
Db        126 GlyAlaIleValMetAlaAlaAlaProAspProTyrValLeuIleSerGlyArgLeuLeu 145

Qy        456 GCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCG 515
Db        146 ValGlyLeuGlyValGlyValAlaSerValThrAlaProValTyrIleAlaGluAlaSer 165

Qy        516 CCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATC 575
Db        166 ProSerGluValArgGlyGlyLeuValSerThrAsnValLeuMetIleThrGlyGlyGln 185

Qy        576 CTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGC 635
Db        186 PheLeuSerTyrLeuValAsnSerAlaPheThrGlnValProGlyThr------TrpArg 203

Qy        636 ATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATG 695
Db        204 TrpMetLeuGlyValSerGlyValProAlaValIleGlnPheIleLeuMetLeuPheMet 223

Qy        696 CCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTG 755
Db        224 ProGluSerProArgTrpLeuPheMetLysAsnArgLysAlaGluAlaIleGlnValLeu 243

Qy        756 GAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCC 815
Db        244 AlaArgThrTyrAspIleSer--------------ArgLeuGluAsp------------ 254

Qy        816 GCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGA 875
Db        255 --------------------GluIleAspHisLeuSerAlaAlaGluGluGluGluLys 267

Qy        876 AACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATC 935
Db        268 GlnArgLysArgThrValGlyTyrLeuAspValPheArgSerLysGluLeuArgLeuAla 287

Qy        936 CTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTC 995
Db        288 PheLeuAlaGlyAlaGlyLeuGlnAlaPheGlnGlnPheThrGlyIleAsnThrValMet 307

Qy        996 TTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGC 1055
Db        308 TyrTyrSerProThrIleValGlnMetAlaGlyPheHisSerAsnGlnLeuAlaLeuPhe 327

Qy       1056 ACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATC 1115
Db        328 LeuSerLeuIleValAlaAlaMetAsnAlaAlaGlyThrValValGlyIleTyrPheIle 347

Qy       1116 GACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTC 1175
Db        348 AspHisCysGlyArgLysLysLeuAlaLeuSerSerLeuPheGlyValIleIleSerLeu 367

Qy       1176 ATC----------------------------------------------------GGCCTCGGC 1187
```

```
                :::                                                    ||||||
Db        368 LeuIleLeuSerValSerPhePheLysGlnSerGluThrSerSerAspGlyGlyLeuTyr 387

Qy       1188 GCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTA 1247
                ||| |||                                          :::||||||
Db        388 GlyTrpLeuAlaVal-----------------------------------LeuGlyLeu 395

Qy       1248 AGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACG 1307
                :::::              |||:::  ||||||:::  |||:::||||||:::
Db        396 AlaLeu--------------TyrIleValPhePheAlaProGlyMetGlyProValPro 410

Qy       1308 TGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGC 1367
                |||      :::|||||||||:::|||  |||  |||  :::        :::
Db        411 TrpThrValAsnSerGluIleTyrProGlnGlnTyrArgGlyIleCysGlyGlyMetSer 430

Qy       1368 GTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAG 1427
                |||  ::::::::  ::::::::  |||||||||:::::::::::
Db        431 AlaThrValAsnTrpIleSerAsnLeuIleValAlaGlnThrPheLeuThrIleAlaGlu 450

Qy       1428 GCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTG 1487
                |||        |||  :::|||  :::  :::|||||||||  ||||||  :::
Db        451 AlaAlaGlyThrGlyMetThrPheLeuIleLeuAlaGlyIleAlaValLeuAlaValIle 470

Qy       1488 TTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTG 1547
                |||        ::::::|||||||||:::|||  |||      |||:::  ::::::
Db        471 PheValIleValPheValProGluThrGlnGlyLeuThrPheSerGluValGluGlnIle 490

Qy       1548 TTCGGCGACACGGCC 1562
                :::  :::  |||
Db        491 TrpLysGluArgAla 495

RESULT 8
E70070
metabolite transport protein homolog ywtG - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
C;Accession: E70070
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: E70070
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-457 <KUN>
A;Cross-references: GB:Z99122; GB:AL009126; NID:g2636029; PIDN:CAB15600.1; PID:g2636109
A;Experimental source: strain 168
C;Genetics:
A;Gene: ywtG
C;Superfamily: glucose transport protein

Alignment Scores:
Pred. No.:                1.27e-31        Length:          457
Score:                    670.50          Matches:         155
Percent Similarity:       53.29%          Conservative:    96
Best Local Similarity:    32.91%          Mismatches:      185
Query Match:              18.55%          Indels:          35
DB:                       2               Gaps:            7

US-10-051-902A-21 (1-2017) x E70070 (1-457)
```

```
Qy        132 AAGAAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCC 191
                ||||||:::  |||:::        |||                    ::::::
Db          2 LysLysGlnSerAsnIleTrpLeuTyrPhe----------PheGlyAlaLeuGlyGly 17

Qy        192 ATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGAC 251
                |||  |||||||||  ||||||:::|||||||||  |||::::::||||||:::
Db         18 AlaLeuTyrGlyTyrAspThrGlyValIleSerGlyAlaIleLeuPheMetLysLysGlu 37

Qy        252 TTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTC 311
                      ::::::          :::::::::      |||   :::   ::::::
Db         38 LeuGlyLeuAsnAlaPheThrGluGlyLeuValValSerSerLeuLeuValGlyAlaIle 57

Qy        312 ATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTG 371
                :::||||||  |||||||||:::  :::|||      ||||||:::      |||:::
Db         58 LeuGlySerGlyAlaAlaGlyLysLeuThrAspArgPheGlyArgLysLysAlaIleMet 77

Qy        372 TTCGCCGCCGTCATATTCTTCGCGGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCC 431
                ||||||::::::|||      |||          :::      |||  |||
Db         78 AlaAlaAlaLeuLeuPheCysIleGlyGlyLeuGlyValAlaLeuAlaProAsnThrGly 97

Qy        432 ATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCG 491
                :::::::::      |||  :::  |||:::  ||||||  :::      |||
Db         98 ValMetValLeuPheArgIleIleLeuGlyLeuAlaValGlyThrSerThrThrIleVal 117

Qy        492 CCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCG 551
                |||:::|||  :::|||::::::|||          ||||||  |||:::|||
Db        118 ProLeuTyrLeuSerGluLeuAlaProLysHisLysArgGlyAlaLeuSerSerLeuAsn 137

Qy        552 GAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGC 611
                ::::::  |||      ||||||||||  |||:::  ||||||  |||:::
Db        138 GlnLeuMetIleThrValGlyIleLeuLeuSerTyrIleValAsnTyrIlePheAlaAsp 157

Qy        612 TTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTG 671
                                  ||||||  |||||||||:::  |||  ||||||:::|||
Db        158 AlaGlu---------AlaTrpArgTrpMetLeuGlyLeuAlaAlaValProSerLeuLeu 174

Qy        672 CTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGC 731
                |||  :::  :::|||  ||||||||||||||||||||  |||
Db        175 LeuLeuIleGlyIleLeuPheMetProGluSerProArgTrpLeuPheThrAsnGlyGlu 194

Qy        732 CTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAG 791
                  :::  ||||||  :::|||||||||      |||  ::::::      :::
Db        195 GluSerLysAlaLysLysIleLeuGluLysLeuArgGlyThr---LysAspIleAspGln 213

Qy        792 CGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTG 851
                  :::  |||||||||  |||
Db        214 GluIleHisAspIleLysGluAla---------------------------------- 221

Qy        852 ACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCC 911
                      |||:::      |||      |||   ::::::
Db        222 ------GluLysGlnAspGluGlyGlyLeuLysGluLeuPhe---------------- 233

Qy        912 CCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCAT 971
                   |||   :::|||      |||::::::|||:::|||:::   |||  |||
Db        234 ---AspProTrpValArgProAlaLeuIleAlaGlyLeuGlyLeuAlaPheLeuGlnGln 252

Qy        972 GCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTA 1031
                   :::|||  :::::::::::::::|||:::|||      |||  :::  |||
Db        253 PheIleGlyThrAsnThrIleIleTyrTyrAlaProLysThrPheThrAsnValGlyPhe 272

Qy       1032 ACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTC 1091
                |||              ||||||||      |||              |||
Db        273 GlyAsnSerAlaSerIleLeuGlyThrVal---GlyIleGlyThrValAsnValLeuMet 291

Qy       1092 ATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGC 1151
                |||:::|||        ||||||   :::||||||:::||||||||||||
Db        292 ThrLeuValAlaIleLysIleIleAspLysIleGlyArgLysProLeuLeuLeuPheGly 311

Qy       1152 ACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAG 1211
                   |||:::::::::|||||||||   |||
Db        312 AsnAlaGlyMetValIleSerLeuIleValLeuAlaLeuValAsnLeuPhePheAsnAsn 331

Qy       1212 CACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTAC 1271
```

```
Db        332 ThrProAlaAlaSerTrpThrThrValIleCysLeuGlyVal--------------Phe 346

Qy       1272 GTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTC 1331
Db        347 IleValValPheAlaValSerTrpGlyProValValTrpValMetLeuProGluLeuPhe 366

Qy       1332 CCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGC 1391
Db        367 ProLeuHisValArgGlyIleGlyThrGlyValSerThrLeuMetLeuHisValGlyThr 386

Qy       1392 GGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTC 1451
Db        387 LeuIleValSerLeuThrTyrProIleLeuMetGluAlaIleGlyIleSerTyrLeuPhe 406

Qy       1452 TTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAG 1511
Db        407 LeuIleTyrAlaAlaIleGlyIleMetAlaPheLeuPheValArgPheLysValThrGlu 426

Qy       1512 ACCCGCGGCCGGACGCTGGAGGAGATGAGCAAG 1544
Db        427 ThrLysGlyArgSerLeuGluGluIleGluGln 437

RESULT 9
F71431
hypothetical protein - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
A;Variety: columbia
C;Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 20-Jun-2000
C;Accession: F71431
R;Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk
 P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenegger, T.; Pohl, T.M.; Terryn, N.; Giel
avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
Nature 391, 485-488, 1998
A;Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech
erhoft, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ans
C.; Chalwatzis, N.
A;Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal
A;Reference number: A71400; MUID:98121113; PMID:9461215
A;Accession: F71431
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-582 <BEV>
A;Cross-references: GB:Z97341; NID:g2244991; PID:g2245004
C;Genetics:
A;Map position: 4COP9-4G3845
C;Superfamily: glucose transport protein

Alignment Scores:
Pred. No.:                3.19e-29        Length:          582
Score:                    629.50          Matches:         166
Percent Similarity:       46.69%          Conservative:    102
Best Local Similarity:    28.92%          Mismatches:      206
Query Match:              17.42%          Indels:          100
DB:                       2               Gaps:            8

US-10-051-902A-21 (1-2017) x F71431 (1-582)

Qy        149 CCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGA 208
Db         25 ProTyrIleMetArgLeu-AlaLeuSerAlaGlyIleGlyGlyLeuLeuPheGlyTyrAs 44

Qy        209 TATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGG 268
Db         44 pThrGlyValIleSerGlyAlaLeuLeuPheIleLysGluAspPheAspGluValAspLy 64

Qy        269 GAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTAC---------TCGCTCATCGGCTC 319
Db         64 sLysThrTrpLeuGlnSerThrIleValSerMetAlaValAlaGlyAlaIleValGlyAl 84

Qy        320 CTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGC 379
Db         84 aAlaValGlyGlyTrpIleAsnAspLysPheGlyArgArgMetSerIleLeuIleAlaAs 104

Qy        380 CGTCATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCAT 439
Db        104 pValLeuPheLeuIleGlyAlaIleValMetAlaPheAlaProAlaProTrpValIleIl 124

Qy        440 GTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTA 499
Db        124 eValGlyArgIlePheValGlyPheGlyValGlyMetAlaSerMetThrSerProLeuTy 144

Qy        500 CACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTT 559
Db        144 rIleSerGluAlaSerProAlaArgIleArgGlyAlaLeuValSerThrAsnGlyLeuLe 164

Qy        560 CATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCT 619
Db        164 uIleThrGlyGlyGlnPhePheSerTyrLeuIleAsnLeuAlaPheValHisThrProGl 184

Qy        620 GAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCT 679
Db        184 yThr------TrpArgTrpMetLeuGlyValAlaGlyValProAlaIleValGlnPheVa 202

Qy        680 CATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGA 739
Db        202 lLeuMetLeuSerLeuProGluSerProArgTrpLeuTyrArgLysAspArgIleAlaGl 222

Qy        740 CGCCAAGGTGGTGCTGGAGAAG------ACCTCCGACACGGCGGAGGAGGCCGCGGAGCG 793
Db        222 uSerArgAlaIleLeuGluArgIleTyrProAlaAspGluValGluAlaGluMetGluAl 242

Qy        794 CCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGAC 853
Db        242 aLeuLysLeuSerValGluAlaGluLysAlaAspGluAlaIleIleGlyAspSerPheSe 262

Qy        854 CGTCCCCAAGAGAGGGAGC---GGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTC 910
Db        262 rAlaLysLeuLysGlyAlaPheGlyAsn------------------------------ 271

Qy        911 CCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCA 970
Db        272 -------ProValValArgArgGlyLeuAlaAlaGlyIleThrValGlnValAlaGlnGl 289

Qy        971 TGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATT 1030
Db        289 nPheValGlyIleAsnThrValMetTyrTyrSerProSerIleValGlnPheAlaGlyTy 309

Qy       1031 AACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTT 1090
Db        309 rAlaSerAsnLysThrAlaMetAlaLeuSerLeuIleThrSerGlyLeuAsnAlaLeuGl 329

Qy       1091 CATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAG 1150
Db        329 ySerIleValSerMetMetPheValAspArgTyrGlyArgArgLysLeuMetIleIleSe 349

Qy       1151 CACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCA 1210
Db        349 rMetPheGlyIleIleAlaCysLeuIleIleLeuAlaThrValPheSerGlnAlaAlaIl 369

Qy       1211 GCAC----------------------------------------CCCGACGCCAAGATA-- 1229
Db        369 eHisAlaProLysIleAspAlaPheGluSerArgThrPheAlaProAsnAlaThrCysSe 389

Qy       1229 ------------------------------------------------------------ 1229
Db        389 rAlaTyrAlaProLeuAlaAlaGluAsnAlaProProSerArgTrpAsnCysMetLysCy 409

Qy       1230 ----------------------------------------CCTTGGGCCATCGGC----- 1244
Db        409 sLeuArgSerGluCysGlyPheCysAlaSerGlyValGlnProTyrAlaProGlyAlaCy 429

Qy       1244 ------------------------------------------------------------ 1244
Db        429 sValValLeuSerAspAspMetLysAlaThrCysSerSerArgGlyArgThrPhePheLy 449
```

```
Qy       1245 ----------------------------CTAAGCATCGCCTCCACCCTCGCCTACGTCGC 1276
                                          |||:::|||            |||:::
Db        449 sAspGlyCysProSerLysPheGlyPheLeuAlaIleValPheLeuGlyLeuTyrIleVa 469

Qy       1277 CTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCT 1336
                  ::::::   |||:::|||   :::    |||:::   :::|||||||||:::|||||
Db        469 lValTyrAlaProGlyMetGlyThrValProTrpIleValAsnSerGluIleTyrProLe 489

Qy       1337 CCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGT 1396
              |:::   |||   |||||      :::         :::|||   |||::::::   ::
Db        489 uArgTyrArgGlyLeuGlyGlyGlyIleAlaAlaValSerAsnTrpValSerAsnLeuIl 509

Qy       1397 CATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCT 1456
               ::::|||   :::||||||||||||:::   |||:::         |||:::|||   ||
Db        509 eValSerGluSerPheLeuSerLeuThrHisAlaLeuGlySerSerGlyThrPheLeuLe 529

Qy       1457 CTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCG 1516
               |::::::|||   :::   :::         |||   :::     :::|||||||||:::
Db        529 uPheAlaGlyPheSerThrIleGlyLeuPhePheIleTrpLeuLeuValProGluThrLy 549

Qy       1517 CGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATC 1576
              :|||         ||||||:::   |||||                   |||
Db        549 sGlyLeuGlnPheGluGluValGluLysLeuLeuGluValGlyPheLysProSerLeuLe 569

Qy       1577 AGACGAGCCAGCCAAGGAGAAGAAGAAGGTGGAAATGGCC 1616
                                |||:::   |||:::|||:::   |||
Db        569 uArgArgArgGluLysLysGlyLysGluValAspAlaAla 582
```

RESULT 10
D86426
hypothetical protein F12P21.2 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 17-May-2002
C;Accession: D86426
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C. C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: D86426
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-580 <STO>
A;Cross-references: GB:AE005172; NID:g11120774; PIDN:AAG30955.1; GSPDB:GN00141
C;Genetics:
A;Map position: 1
C;Superfamily: glucose transport protein

```
Alignment Scores:
Pred. No.:              5.1e-29      Length:         580
Score:                  626.00       Matches:        159
Percent Similarity:     47.83%       Conservative:   106
Best Local Similarity:  28.70%       Mismatches:     182
Query Match:            17.32%       Indels:         107
DB:                     2            Gaps:           11

US-10-051-902A-21 (1-2017) x D86426 (1-580)

Qy        143 CAACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGG 202
               :::   |||    |||||||| |||         |||   :::      :::|||   ||
Db         24 LysAsnProTyrValLeuArgLeu-AlaPheSerAlaGlyIleGlyGlyLeuLeuPheGl 43

Qy        203 CTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAG 262
              ||||||   ||||||:::|||||||||   |||||||||:::   |||||
Db         43 yTyrAspThrGlyValIleSerGlyAlaLeuLeuTyrIleArgAspAspPheLysSerVa 63

Qy        263 TGACGGGAAGGTG---------GAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCAT 313
               |||                   |||:::::::::   :::              :::::::
Db         63 lAspArgAsnThrTrpLeuGlnGluMetIleValSerMetAlaValAlaGlyAlaIleVa 83

Qy        314 CGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTT 373
              :|||:::         |||      :::|||   :::|||||||||     |||:::
Db         83 lGlyAlaAlaIleGlyGlyTrpAlaAsnAspLysLeuGlyArgArgSerAlaIleLeuMe 103

Qy        374 CGCCGCCGTCATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCAT 433
               |||      :::|||      |||      :::|||         |||   |||   ::::
Db        103 tAlaAspPheLeuPheLeuLeuGlyAlaIleIleMetAlaAlaAlaProAsnProSerLe 123

Qy        434 GCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCC 493
              :|||:::   |||||      |||:::|||||||||   |||   |||   |||||
Db        123 uLeuValValGlyArgValPheValGlyLeuGlyValGlyMetAlaSerMetThrAlaPr 143

Qy        494 GGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGA 553
              |:::|||   :::|||   |||||||||         ||||||   |||   |||
Db        143 oLeuTyrIleSerGluAlaSerProAlaLysIleArgGlyAlaLeuValSerThrAsnGl 163

Qy        554 GGTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTT 613
                 |||      |||      |||      |||:::      |||   ||||||:::
Db        163 yPheLeuIleThrGlyGlyGlnPheLeuSerTyrLeuIleAsnLeuAlaPheThr----- 181

Qy        614 GCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCT 673
                 :::         ||||||   ||||||||||||         |||::::::|||
Db        182 -AspValThrGlyThrTrpArgTrpMetLeuGlyIleAlaGlyIleProAlaLeuLeuGl 201

Qy        674 CGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCT 733
                 :::::::::      :::||||||||||||||||||||         |||||||||
Db        201 nPheValLeuMetPheThrLeuProGluSerProArgTrpLeuTyrArgLysGlyArgGl 221

Qy        734 CGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCG 793
                 :::||||||   :::|||   :::         :::||||||:::      :::
Db        221 uGluGluAlaLysAlaIleLeuArgArgIleTyr---SerAlaGluAspValGluGlnGl 240

Qy        794 CCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGAC 853
               :::      :::|||                                    |||
Db        240 uIleArgAlaLeuLys---------------------------AspSerValGluThrGl 251

Qy        854 CGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCC 913
               :::   :::   ||||||      :::      :::   |||         |||
Db        251 uIleLeuGluGluGlySerSerGluLysIleAsnMetIleLys---------LeuCysLy 268

Qy        914 GACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGC 973
                    :::||||||   |||::::::|||:::|||:::         ||||||
Db        268 sAlaLysThrValArgArgGlyLeuIleAlaGlyValGlyLeuGlnValPheGlnGlnPh 288

Qy        974 GTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAAC 1033
               :::||||||::::::|||::::::|||||||||   :::      :::      |||
Db        288 eValGlyIleAsnThrValMetTyrTyrSerProThrIleValGlnLeuAlaGlyPheAl 308

Qy       1034 GAACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCAT 1093
               :::::::::         |||         :::
Db        308 aSerAsnArgThrAlaLeuLeuLeuSerLeuValThrAlaGlyLeuAsnAlaPheGlySe 328

Qy       1094 CTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCAC 1153
               :::::::::   :::|||||||||   :::|||||||:::   ||||||:::   |||
Db        328 rIleIleSerIleTyrPheIleAspArgIleGlyArgLysLysLeuLeuIleIleSerLe 348

Qy       1154 GGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTC---------------- 1196
                 |||:::|||:::||||||   |||:::      |||:::
Db        348 uPheGlyValIleIleSerLeu---GlyIleLeuThrGlyValPheTyrGluAlaAlaTh 367

Qy       1196 ------------------------------------------------------------ 1196

Db        367 rHisAlaProAlaIleSerSerLeuGluThrGlnArgPheAsnAsnIleSerCysProAs 387

Qy       1196 ------------------------------------------------------------ 1196

Db        387 pTyrLysSerAlaMetAsnThrAsnAlaTrpAspCysMetThrCysLeuLysAlaSerSe 407

Qy       1197 ---------------------ACCGTCGTCGGC---CAGCACCCCGACGCC-------- 1223
                                   :::   :::|||   :::||||||   |||
```

```
Db        407 rProSerCysGlyTyrCysSerSerProIleGlyLysGluHisProGlyAlaCysTrpIl 427

Qy       1223 ------------------------------------------------------------ 1223

Db        427 eSerAspAspSerValLysAspLeuCysHisAsnGluAsnArgLeuTrpTyrThrArgGl 447

Qy       1224 ----------AAGATACCTTGG------GCCATCGGCCTAAGCATCGCCTCCACCCTCGC 1267
                            |||         :::||||||   :::
Db        447 yCysProSerAsnPheGlyTrpPheAlaLeuLeuGlyLeuGlyLeu-------------- 462

Qy       1268 CTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGAT 1327
               |||:::   |||||||||   |||:::|||   :::   |||:::   :::|||||||||
Db        463 -TyrIleIlePhePheSerProGlyMetGlyThrValProTrpIleValAsnSerGluIl 482

Qy       1328 CTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCAC 1387
               |:::||||||:::   |||   :::         :::         ||||||   :::::
Db        482 eTyrProLeuArgPheArgGlyIleCysGlyGlyIleAlaAlaThrAlaAsnTrpIleSe 502

Qy       1388 CAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAG 1447
               ::::   :::::::::   :::||||||||||||:::::||||||          ::
Db        502 rAsnLeuIleValAlaGlnSerPheLeuSerLeuThrGluAlaIleGlyThrSerTrpTh 522

Qy       1448 CTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCC 1507
               :|||   ::::::      |||:::   :::|||   :::|||            :::||
Db        522 rPheLeuIlePheGlyValIleSerValIleAlaLeuLeuPheValMetValCysValPr 542

Qy       1508 GGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTG 1547
               |||||||:::|||      :::||||||:::   |||:::
Db        542 oGluThrLysGlyMetProMetGluGluIleGluLysMet 555

RESULT 11
D70073
metabolite transport protein homolog yxcC - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
C;Accession: D70073
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: D70073
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-461 <KUN>
A;Cross-references: GB:Z99124; GB:AL009126; NID:g2636442; PIDN:CAB16017.1; PID:g2636527
A;Experimental source: strain 168
C;Genetics:
A;Gene: yxcC
C;Superfamily: glucose transport protein

Alignment Scores:
Pred. No.:                4.57e-28        Length:          461
Score:                    609.50          Matches:         149
Percent Similarity:       49.59%          Conservative:    93
Best Local Similarity:    30.53%          Mismatches:      195
Query Match:              16.86%          Indels:          51
DB:                       2               Gaps:            11

US-10-051-902A-21 (1-2017) x D70073 (1-461)

Qy        180 TCCATGACCTCCATCCTCCTCGGCTACGATATCGGGGGTGATGAGCGGGGCGTCGCTGTAC 239
```

```
                  ::::::      :::|||   |||||||||   ||||||:::|||||||||   |||:::
Db         15 AlaLeuGlyGlyLeuLeuTyrGlyTyrAspThrGlyValIleSerGlyAlaLeuLeuPhe 34

Qy        240 ATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAAC 299
                |||      |||      ::::::              :::::::::   :::|||
Db         35 IleAsnAsnAspIleProLeuThrThrLeuThrGluGlyLeuValValSerMetLeuLeu 54

Qy        300 CTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGG 359
                |||   ::::::   ||||||      :::|||      ||||||   |||||||||
Db         55 LeuGlyAlaIlePheGlySerAlaLeuSerGlyThrCysSerAspArgTrpGlyArgArg 74

Qy        360 TACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCC 419
                   :::          ::::::||||||      |||                |||:::
Db         75 LysValValPheValLeuSerIleIlePheIleIleGlyAlaLeuAlaCysAlaPheSer 94

Qy        420 GTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCG 479
                        ||||||:::      |||   :::   |||:::   ||||||   :::
Db         95 GlnThrIleGlyMetLeuIleAlaSerArgValIleLeuGlyLeuAlaValGlyGlySer 114

Qy        480 CTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTG 539
                   :::   |||||||||   :::|||::::::|||           ||||||   |||
Db        115 ThrAlaLeuValProValTyrLeuSerGluMetAlaProThrLysIleArgGlyThrLeu 134

Qy        540 ACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTAT 599
                   :::         :::   |||          ||||||||||||   |||:::   ||||||
Db        135 GlyThrMetAsnAsnLeuMetIleValThrGlyIleLeuLeuAlaTyrIleValAsnTyr 154

Qy        600 GCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCG 659
                   |||:::      |||            ||||||   |||:::|||:::   |||
Db        155 LeuPheThr------ProPheGlu---AlaTrpArgTrpMetValGlyLeuAlaAlaVal 171

Qy        660 CCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTC 719
                |||:::|||||||||   :::   :::      ||||||||||||||||||||||||||
Db        172 ProAlaValLeuLeuLeuIleGlyIleAlaPheMetProGluSerProArgTrpLeuVal 191

Qy        720 ATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAG 779
                   :::|||      :::|||:::   ::::::      |||   |||
Db        192 LysArgGlySerGluGluGluAlaArgArgIleMetAsnIleThrHisAsp--------- 208

Qy        780 GAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGAC 839
                                                          |||:::::::::::
Db        209 ----------------------------------------------ProLysAspIleGlu 213

Qy        840 GGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAG 899
                   ::::::      :::   |||:::|||        :::             ||||||
Db        214 MetGluLeuAlaGluMet---LysGlnGlyGluAlaGluLys-----------LysGlu 228

Qy        900 CTCATCCTGTCCCCGACCCCGGCC------ATGCGGCGCATCCTGCTGTCCGGGATCGGC 953
                   |||          |||      :::|||   :::||||||   |||:::|||
Db        229 ThrThrLeuGlyValLeuLysAlaLysTrpIleArgProMetLeuLeuIleGlyValGly 248

Qy        954 ATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTG 1013
                   :::      ||||||   |||:::||||||::::::|||::::::|||:::|||   :::
Db        249 LeuAlaIlePheGlnGlnAlaValGlyIleAsnThrValIleTyrTyrAlaProThrIle 268

Qy       1014 TTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGT 1073
                |||          ||||||                ||||||   |||          |||
Db        269 PheThrLysAlaGlyLeuGlyThrSerAlaSerAlaLeuGly---ThrMetGlyIleGly 287

Qy       1074 GTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGG 1133
                   :::         :::         :::   |||          ||||||   |||||||||:::
Db        288 IleLeuAsnValIleMetCysIleThrAlaMetIleLeuIleAspArgValGlyArgLys 307

Qy       1134 CCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGG 1193
                   ||||||:::      :::   ||||||   |||||||||      |||
Db        308 LysLeuLeuIleTrpGlySerValGlyIleThrLeuSerLeuAlaAlaLeuSerGlyVal 327

Qy       1194 CTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATC 1253
                |||      :::|||                |||         |||         :::::::::
Db        328 LeuLeuThrLeuGlyLeu------SerAlaSerThrAlaTrp---------MetThrVal 342

Qy       1254 GCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTG 1313
                            |||:::   |||:::            ||||||:::   ||||||
```

```
Db        343 ValPheLeuGlyValTyrIleValPheTyrGlnAlaThrTrpGlyProValValTrpVal 362

Qy       1314 TACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCC 1373
                 |||:::||||||   :::   |||
Db        363 LeuMetProGluLeuPheProSerLysAlaArgGlyAlaAlaThrGlyPheThrThrLeu 382

Qy       1374 GCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATC 1433
                     :::   ::::::|||:::   |||      :::      |||:::
Db        383 ValLeuSerAlaAlaAsnLeuIleValSerLeuValPheProLeuMetLeuSerAlaMet 402

Qy       1434 ACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTC 1493
                |||         |||   ::::::|||   |||      |||::::::   |||
Db        403 GlyIleAlaTrpValPheMetValPheSerValIleCysLeuLeuSerPhePhePheAla 422

Qy       1494 TACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAG-----------ATGAGC 1541
                :::    :::||||||||:::|||::::::|||||||||           :::
Db        423 PheTyrMetValProGluThrLysGlyLysSerLeuGluGluIleGluAlaSerLeuLys 442

Qy       1542 AAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAGCCAAGGAGAAGAAG 1601
                |||   |||                                   |||:::||||||
Db        443 LysArgPhe-----------------------------------LysLysLysLys 449

Qy       1602 AAGGTGGAAATGGCCGCCACTAAC 1625
                  :::          |||
Db        450 SerThrGlnAsnGlnValLeuAsn 457
```

RESULT 12
G69789
sugar transporter homolog ydjK - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
C;Accession: G69789
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: G69789
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-473 <KUN>
A;Cross-references: GB:Z99107; GB:AL009126; NID:g2632866; PIDN:CAB12442.1; PID:g2632936
A;Experimental source: strain 168
C;Genetics:
A;Gene: ydjK
C;Superfamily: glucose transport protein

Alignment Scores:

Pred. No.:	4.58e-28	Length:	473
Score:	609.50	Matches:	157
Percent Similarity:	52.48%	Conservative:	97
Best Local Similarity:	32.44%	Mismatches:	187
Query Match:	16.86%	Indels:	43
DB:	2	Gaps:	11

US-10-051-902A-21 (1-2017) x G69789 (1-473)

```
Qy        135 AAGAAGGGCAAC---GTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCC 191
                 |||:::||||||   :::   |||                :::::::::::
Db          3 LysGlnGlyAsnGlnMetSerPheLeuArgThrIleIleLeuValSerThrPheGlyGly 22

Qy        192 ATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAG--- 248
                :::   :::::   |||:::   ::::::||||||           :::
Db         23 LeuLeuPheGlyTyrAspThrGlyValLeuAsnGlyAlaLeuProTyrMetGlyGluPro 42

Qy        249 ---GACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTAC 305
                    |||::::::          ::::::         |||
Db         43 AspGlnLeuAsnLeuAsnAlaPheThrGluGlyLeuValThrSerSerLeuLeuPheGly 62

Qy        306 TCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACC 365
                :::   :::|||:::          ||||||   ||||||:::   |||||||||
Db         63 AlaAlaLeuGlyAlaValPheGlyGlyArgMetSerAspPheAsnGlyArgArgLysAsn 82

Qy        366 ATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGGSGTTCCTCATGGGGTTCGCCGTCAAC 425
                |||:::|||   |||||||||||||||                  ||||||   |||
Db         83 IleLeuPheLeuAlaValIlePhePheIleSerThrIleGlyCysThrPheAlaProAsn 102

Qy        426 TACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATG 485
                      :::::::::      |||||||||   ||||||   ||||||   |||   :::
Db        103 ValThrValMetIleIleSerArgPheValLeuGlyIleAlaValGlyGlyAlaSerVal 122

Qy        486 ATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCG 545
                      |||   |||   ||||||:::||||||      :::||||||   :::   :::
Db        123 ThrValProAlaTyrLeuAlaGluMetSerProValGluSerArgGlyArgMetValThr 142

Qy        546 TTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTC 605
                      |||:::   |||      |||   ||||||   :::|||   |||
Db        143 GlnAsnGluLeuMetIleValSerGlyGlnLeuLeuAlaPheValPheAsn--------- 159

Qy        606 TCCCGCTTGCCGCTGAACCTCGGG--------------TGGCGCATCATGCTCGGCATC 650
                :::   |||          :::|||                ||||||   ||||||   |||
Db        160 AlaIleLeuGlyThrThrMetGlyAspAsnSerHisValTrpArgPheMetLeuValIle 179

Qy        651 GGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGG 710
                  :::   |||::::::   |||          ::::::   ||||||||||||||||||
Db        180 AlaSerLeuProAlaLeuPheLeuPhePheGlyMetIleArgMetProGluSerProArg 199

Qy        711 TGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGAC 770
                |||||||||   |||||||||      ||||||      ||||||:::|||      |||
Db        200 TrpLeuValSerLysGlyArgLysGluAspAlaLeuArgValLeuLysLysIleArgAsp 219

Qy        771 ACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAG 830
                      |||||||||   |||   :::|||:::   |||            :::
Db        220 GluLysArgAlaAlaAlaGlu---LeuGlnGluIleGluPheAlaPheLysLysGluAsp 238

Qy        831 GAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTG 890
                :::|||                                          ||||||
Db        239 GlnLeu-----------------------------------------GluLysAlaThr 244

Qy        891 TGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATC 950
                :::|||:::|||              |||   :::|||||||||:::      |||:::
Db        245 PheLysAspLeu---------SerValProTrpValArgArgIleValPheIleGlyLeu 261

Qy        951 GGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTC 1010
                ||||||         |||         |||::::::|||:::::::::|||
Db        262 GlyIleAlaIleValGlnGlnIleThrGlyValAsnSerIleMetTyrTyrGlyThrGlu 281

Qy       1011 GTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTC 1070
                :::   ::::::   |||         :::         :::|||
Db        282 IleLeuArgAsnSerGlyPheGlnThrGluAlaAlaLeuIleGly--------------- 296

Qy       1071 GGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTC-----------TTCATCGAC 1118
                   :::         :::      :::||||||||||||                :::
Db        297 AsnIleAlaAsnGlyValIleSerValLeuAlaThrPheValGlyIleTrpLeuLeuGly 316

Qy       1119 GGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATC 1178
                   ||||||||||||||:::|||:::                |||         :::|||:::
Db        317 ArgValGlyArgArgProMetLeuMetThrGlyLeuIleGlyThrThrThrAlaLeuLeu 336

Qy       1179 GGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCC 1238
                   :::|||         |||:::      |||      :::|||:::
Db        337 LeuIleGlyIlePheSerLeuValLeuGluGlySerPro------AlaLeuProTyrVal 354

Qy       1239 ATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGC 1298
                :::         ::::::   |||:::   ::::::||||||            :::
```

```
Db        355 Val---------LeuSerLeuThrValThrPheLeuAlaPheGlnGlnGlyAlaIleSer 371

Qy       1299 CCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGC 1358
              |||:::|||||:::      ||||||||||||||||||:::::|||   ||||||
Db        372 ProValThrTrpLeuMetLeuSerGluIleProLeuArgLeuArgGlyLeuGlyMet 391

Qy       1359 TCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCG 1418
                 :::   |||           :::   :::      :::|||   ||||||
Db        392 GlyValThrValPheCysLeuTrpMetValAsnPheAlaValSerPheThrPheProIle 411

Qy       1419 CTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTC 1478
              |||      ||||||   :::      :::|||||:::::   |||:::
Db        412 LeuLeuAlaAlaIleGlyLeuSerThrThrPhePheIlePheValGlyLeuGlyIleCys 431

Qy       1479 GCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATG 1538
              :::   :::|||        :::||||||||||||:::|||   :::||||||::::::
Db        432 SerValLeuPheValLysArgPheLeuProGluThrLysGlyLeuSerLeuGluGlnLeu 451

Qy       1539 AGCAAGCTGTTC 1550
                 :::   |||
Db        452 GluGluAsnPhe 455
```

RESULT 13
T38125
myo-inositol transporter 2 - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
C;Accession: T38125
R;Badcock, K.; Churcher, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, May 1997
A;Reference number: Z21772
A;Accession: T38125
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-557 <BAD>
A;Cross-references: EMBL:Z95334; PIDN:CAB08597.1; GSPDB:GN00066; SPDB:SPAC20G8.03
A;Experimental source: strain 972h-; cosmid c20G8
C;Genetics:
A;Gene: SPDB:SPAC20G8.03
A;Map position: 1
C;Superfamily: maltose transport protein MAL61

Alignment Scores:

Pred. No.:	1.46e-27	Length:	557
Score:	601.00	Matches:	160
Percent Similarity:	51.22%	Conservative:	112
Best Local Similarity:	30.13%	Mismatches:	211
Query Match:	16.63%	Indels:	48
DB:	2	Gaps:	16

US-10-051-902A-21 (1-2017) x T38125 (1-557)

```
Qy         93 GAGATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGG 152
              |||   |||:::|||      |||:::|||                  :::
Db         52 GluGluAlaAsnAlaThrAspProGlnAlaAsnGluValAlaAspGluAsnGlyGluGly 71

Qy        153 TTC-----------------------GCCTTCGCCTGCGCCATCCTCGCCTCCATGACC 188
              |||                        :::      :::   :::|||   ::::::
Db         72 PheGluAlaGluLysIleSerSerTrpIleTrpValLeuSerAlaValAlaGlyIleSer 91

Qy        189 TCCATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAG 248
              :::|||   |||||||||   ||||||:::|||||||||        :::
Db         92 GlyLeuLeuPheGlyTyrAspThrGlyValIleSerGlyAlaLeuAlaValLeuGlySer 111

Qy        249 GACTTCAAC------ATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTC 302
              |||            :::|||   |||:::   |||::::::          :::
Db        112 AspLeuGlyHisValLeuSerSerGlyGlnLysGluLeuIleThrSerAlaThrSerPhe 131

Qy        303 TACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTAC 362
                 :::||||||   :::      :::|||      :::||||||:::||||||:::
Db        132 AlaAlaLeuIleSerAlaThrThrSerGlyTrpLeuAlaAspTrpValGlyArgLysArg 151

Qy        363 ACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGGSGTTCCTCATGGGGTTCGCCGTC 422
```

```
                 ::::::   |||      ||||||      |||      :::|||      :::
Db        152 LeuLeuLeuCysAlaAspAlaIlePheValIleGlySerValIleMetAlaAlaSerArg 171

Qy        423 AACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTC 482
              |||   ||||||::::::   ||||||||:::   |||   |||:::|||
Db        172 AsnValAlaMetMetValValGlyArgPheIleValGlyTyrGlyIleGlyLeuThrSer 191

Qy        483 ATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACG 542
              :::|||   |||:::|||      |||::::::||||||      ||||||   |||
Db        192 LeuIleValProMetTyrIleThrGluLeuAlaProAlaArgLeuArgGlyArgLeuVal 211

Qy        543 TCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCT 602
                          |||||||||      |||   |||:::   |||      |||   |||
Db        212 IleIleTyrValValPheIleThrGlyGlyGlnLeuIleAlaTyrSerLeuAsnAlaAla 231

Qy        603 TTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCG 662
              |||              ::::::   ||||||||||||||   ||||||||||||||||
Db        232 PheGluHis------ValHisGlnGlyTrpArgIleMetPheGlyIleGlyAlaAlaPro 249

Qy        663 TCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATG 722
              ::::::          :::   :::          |||||||||||||||:::|||:::
Db        250 AlaLeuGlyGlnLeuIleSerLeuPheTrpThrProGluSerProArgTyrLeuLeuArg 269

Qy        723 AAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAG--- 779
                    :::                 :::|||   :::                |||:::
Db        270 HisAsnHisValGluLysValTyrLysIleLeuSerArgIleHisProGluAlaLysPro 289

Qy        780 ---GAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTC 836
                 |||   |||   :::::::::   |||:::
Db        290 AlaGluIleAlaTyrLysValSerLeuIleGln---------------------------- 300

Qy        837 GACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAG---AAGCGGGTGTGG 893
              :::|||      |||      |||:::      ||||||:::   :::      :::
Db        301 GluGlyValLysValAspPheProGlu---------GlyAsnLysPheGlnHisPhePhe 317

Qy        894 AAGGAGCTC---ATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATC 950
                    |||   :::|||          |||:::   ||||||   |||      |||
Db        318 HisSerLeuLysValLeuPheThrValProSerAsnArgArgSerLeuPheIleGlyCys 337

Qy        951 GGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTC 1010
                 :::   :::||||||          |||   :::::::::   ::::::|||   :::
Db        338 PheLeuGlnTrpPheGlnGlnPheSerGlyThrAsnAlaIleGlnTyrPheSerAlaIle 357

Qy       1011 GTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTC 1070
              :::|||:::|||   |||      |||          :::      :::
Db        358 IlePheGlnSerValGlyPheLysAsnSer---------IleSerValSerIleValVal 374

Qy       1071 GGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGG 1130
              |||   |||      :::|||   ::::::|||      |||||||||   :::||||||
Db        375 GlyAlaThrAsnPheValPheThrIleValAlaPheMetPheIleAspArgIleGlyArg 394

Qy       1131 CGGCCGCTGTTGCTGGGCAGCACGGGCGGGGATAATCCTCTCCCTCATCGGCCTCGGCGCC 1190
              |||   :::||||||      |||      ::::::::::::           |||
Db        395 ArgArgIleLeuLeu---CysThrSerAlaValMetIleAla-----------GlyLeu 409

Qy       1191 GGGCTCACCGTCGTCGGCCAGCAC--------------------CCCGACGCCAAGATA 1229
                 |||      :::      |||                        :::
Db        410 AlaLeuCysAlaIleAlaTyrHisPheLeuProAlaAspThrThrGlnAsnThrAsnSer 429

Qy       1230 CCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATC 1289
              |||      :::   :::||||||   :::   :::::::|||   ::::::
Db        430 GlyTrpGlnTyr---ValValLeuAlaSerIleIleIlePheLeuAlaSerTyrAlaSer 448

Qy       1290 GGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCG 1349
              |||:::|||   |||   |||          :::|||:::||||||:::::|||||||||
Db        449 GlyIleGlyAsnIleProTrp---GlnGlnAlaGluLeuPheProMetGluValArgAla 467

Qy       1350 CTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACC 1409
              ||||||                |||   |||   |||   :::   :::||||||   :::
Db        468 LeuGlyAlaGlyPheSerThrAlaIleAsnTrpValGlyAsnLeuIleIleSerAlaSer 487

Qy       1410 TTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATC 1469
              ||||||::::::   :::::::||||||      |||:::|||   |||:::::::|||
```

```
Db        488 PheLeuThrMetMetGluSerIleThrProThrGlyThrPheAlaLeuPheAlaGlyPhe 507

Qy       1470 GCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTG 1529
                 :::      |||      |||          ||||||      |||   ::::::
Db        508 CysPheValGlyLeuValThrSerTyrPheThrTyrProGluLeuAlaGlyMetSerIle 527

Qy       1530 GAGGAGATGAGCAAGCTGTTC------GGCGACACGGCCGCCGCCTCGGAATCAGACGAG 1583
              |||   :::   ||||||          |||          |||      ||||||   :::
Db        528 GluAsnIleHisLysLeuLeuGluLysGlyPheTrpGlnAlaValLysGluSerThrLys 547

Qy       1584 CCAGCCAAGGAGAAGAAGAAGGTGGAAATGGCC 1616
                 :::   |||   :::::::::   |||
Db        548 ArgValArg---LysGlyArgIleAspGluAla 557

RESULT 14
D84772
probable sugar transporter [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 17-May-2002
C;Accession: D84772
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: D84772
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-580 <STO>
A;Cross-references: GB:AE002093; NID:g4263781; PIDN:AAD15441.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g35740
A;Map position: 2
C;Superfamily: glucose transport protein

Alignment Scores:
Pred. No.:                1.02e-26      Length:          580
Score:                    586.50        Matches:         154
Percent Similarity:       46.38%        Conservative:    102
Best Local Similarity:    27.90%        Mismatches:      198
Query Match:              16.23%        Indels:          98
DB:                       2             Gaps:            8

US-10-051-902A-21 (1-2017) x D84772 (1-580)

Qy        143 CAACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGG 202
                 :::   |||      :::||||||  |||:::   |||   :::      :::|||   ||
Db         22 GluThrProTyrIleMetArgLeu-AlaLeuSerAlaGlyIleGlyGlyLeuLeuPheGl 41

Qy        203 CTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAG 262
                 ||||:::   ||||||::::::||||||   |||||||||||||:::::|||
Db         41 yTyrAsnThrGlyValIleAlaGlyAlaLeuLeuTyrIleLysGluGluPheGlyGluVa 61

Qy        263 TGACGGGAAGGTG---------GAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCAT 313
                 |||   |||              |||:::::::::   :::            ::::::::
Db         61 lAspAsnLysThrTrpLeuGlnGluIleIleValSerMetThrValAlaGlyAlaIleVa 81

Qy        314 CGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTT 373
                 :|||:::          |||      :::|||      ||||||||   :::::::::
Db         81 lGlyAlaAlaIleGlyGlyTrpTyrAsnAspLysPheGlyArgArgMetSerValLeuIl 101

Qy        374 CGCCGCCGTCATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCAT 433
                 |||   |||:::|||      |||      :::|||      |||            ::
Db        101 eAlaAspValLeuPheLeuGlyAlaLeuValMetValIleAlaHisAlaProTrpVa 121

Qy        434 GCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCC 493
                 :::::::   ||||||   :::   |||   |||||||||   |||   |||   :::||
Db        121 lIleIleLeuGlyArgLeuLeuValGlyPheGlyValGlyMetAlaSerMetThrSerPr 141

Qy        494 GGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGA 553
                 |:::|||   :::|||:::|||||||||      ||||||   |||   |||
Db        141 oLeuTyrIleSerGluMetSerProAlaArgIleArgGlyAlaLeuValSerThrAsnGl 161

Qy        554 GGTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTT 613
                 :::   |||      |||      |||   |||:::   |||   ||||||
Db        161 yLeuLeuIleThrGlyGlyGlnPheLeuSerTyrLeuIleAsnLeuAlaPheValHisTh 181

Qy        614 GCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCT 673
                 |||            ||||||   |||||||||:::   |||   |||::::::::
Db        181 rProGlyThr------TrpArgTrpMetLeuGlyValSerAlaIleProAlaIleIleGl 199

Qy        674 CGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCT 733
                    ::::::|||   :::||||||||||||||||||||||           |||
Db        199 nPheCysLeuMetLeuThrLeuProGluSerProArgTrpLeuTyrArgAsnAspArgLy 219

Qy        734 CGCGGACGCCAAGGTGGTGCTGGAGAAG------ACCTCCGACACGGCGGAGGAGGCCGC 787
                 |||:::::::::   :::||||||:::      ::::::     |||
Db        219 sAlaGluSerArgAspIleLeuGluArgIleTyrProAlaGluMetValGluAlaGluIl 239

Qy        788 GGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGT 847
                    |||   :::         |||            |||::::::   |||
Db        239 eAlaAlaLeuLysGluSerValArgAlaGluThrAlaAspGluAspIleIleGlyHisTh 259

Qy        848 GGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCT 907
                    :::         ||||||:::                              ||
Db        259 rPheSerAspLysLeuArgGlyAla--------------------------------Le 268

Qy        908 GTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCA 967
                 ||||      |||   :::|||      |||   :::||||||   :::      ||
Db        268 uSer---AsnProValValArgHisGlyLeuAlaAlaGlyIleThrValGlnValAlaGl 287

Qy        968 GCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGG 1027
                 |      :::||||||:::::|||:::::|||||||||   :::   :::      ||
Db        287 nGlnPheValGlyIleAsnThrValMetTyrTyrSerProThrIleLeuGlnPheAlaGl 307

Qy       1028 ATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCT 1087
                 |      ::::::|||      :::                              ::
Db        307 yTyrAlaSerAsnLysThrAlaMetAlaLeuAlaLeuIleThrSerGlyLeuAsnAlaVa 327

Qy       1088 TTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGG 1147
                 :      :::::::::      |||:::|||      |||||||||   |||::::::
Db        327 lGlySerValValSerMetMetPheValAspArgTyrGlyArgArgLysLeuMetIleIl 347

Qy       1148 CAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGG 1207
                 |||      |||||||||      |||:::   |||   |||
Db        347 eSerMetPheGlyIleIleThrCysLeuValIleLeuAlaAlaValPheAsnGluAlaSe 367

Qy       1208 CCAGCAC---------------------------------------------------- 1214
                    |||
Db        367 rAsnHisAlaProLysIleAspLysArgAspSerArgAsnPheAlaLysAsnAlaThrCy 387

Qy       1214 ---------------------------------------------------------- 1214

Db        387 sProAlaPheAlaProPheThrAlaSerArgSerProProSerAsnTrpAsnCysMetLy 407

Qy       1215 ----------------------------------------------------------CCCGACGC 1222
                                                                        |||   ||
Db        407 sCysLeuGlnTyrAspCysGlyPheCysSerAsnGlyAlaGlnGluTyrAlaProGlyAl 427

Qy       1223 CAAGATA---------------------------------------------------- 1229
                 |   |||
Db        427 aCysIleValGlnSerAlaAspMetLysAlaLeuCysHisSerLysGlyArgThrPhePh 447

Qy       1230 -------------CCTTGGGCCATCGGC---CTAAGCATCGCCTCCACCCTCGCCTACGT 1273
                          |||          |||   |||:::|||                |||::
Db        447 eLysAspGlyCysProSerLysPheGlyTyrLeuAlaIleValPheLeuGlyLeuTyrIl 467

Qy       1274 CGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCC 1333
                 :      ::::::   |||:::|||   :::   |||:::   :::||||||||:::||
Db        467 eIleValTyrAlaProGlyMetGlyThrValProTrpIleValAsnSerGluIleTyrPr 487

Qy       1334 GCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGG 1393
                 ||||:::   |||   |||          :::          :::|||   :::::::::
Db        487 oLeuArgTyrArgGlyLeuAlaGlyGlyIleAlaAlaValSerAsnTrpMetSerAsnLe 507
```

```
Qy       1394 CGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTT 1453
Db        507 uValValSerGluThrPheLeuThrLeuThrAsnAlaValGlySerSerGlyThrPheLe 527

Qy       1454 CCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGAC 1513
Db        527 uLeuPheAlaGlySerSerAlaValGlyLeuPhePheIleTrpLeuLeuValProGluTh 547

Qy       1514 CCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTG 1547
Db        547 rLysGlyLeuGlnPheGluGluValGluLysLeu 558
```

RESULT 15
S47089
arabinose-proton symporter - Klebsiella oxytoca
C;Species: Klebsiella oxytoca
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 24-Sep-1999
C;Accession: S47089
R;Shatwell, K.P.; Charalambous, B.M.; McDonald, T.P.; Henderson, P.J.F.
submitted to the EMBL Data Library, June 1994
A;Description: The nucleotide sequence of the gene araE for arabinose-proton synport in
A;Reference number: S47089
A;Accession: S47089
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-472 <SHA>
A;Cross-references: EMBL:X79598; NID:g498919; PIDN:CAA56110.1; PID:g498920
C;Superfamily: glucose transport protein
C;Keywords: arabinose transport

Alignment Scores:

Pred. No.:	1.31e-26	Length:	472
Score:	584.50	Matches:	143
Percent Similarity:	52.25%	Conservative:	101
Best Local Similarity:	30.62%	Mismatches:	182
Query Match:	16.17%	Indels:	41
DB:	2	Gaps:	10

US-10-051-902A-21 (1-2017) x S47089 (1-472)

```
Qy        168 GCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGG 227
Db         26 SerIleAlaAlaAlaValAlaGlyLeuLeuPheGlyLeuAspIleGlyValIleAlaGly 45

Qy        228 GCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATG 287
Db         46 AlaLeuProPheIleThrAspHisPheValLeuSerSerArgLeuGlnGluTrpValVal 65

Qy        288 GGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGG 347
Db         66 SerSerMetMetLeuGlyAlaAlaIleGlyAlaLeuPheAsnGlyTrpLeuSerPheArg 85

Qy        348 ATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGSGTTCCTC 407
Db         86 LeuGlyArgLysTyrSerLeuMetValGlyAlaValLeuPheValAlaGlySerValGly 105

Qy        408 ATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGC 467
Db        106 SerAlaPheAlaThrSerValGluMetLeuLeuValAlaArgIleValLeuGlyValAla 125

Qy        468 GTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCG 527
Db        126 ValGlyIleAlaSerTyrThrAlaProLeuTyrLeuSerGluMetAlaSerGluAsnVal 145

Qy        528 CGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTAC 587
Db        146 ArgGlyLysMetIleSerMetTyrGlnLeuMetValThrLeuGlyIleValMetAlaPhe 165

Qy        588 GTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGC 647
Db        166 LeuSerAspThrAlaPheSer---------TyrSerGlyAsnTrpArgAlaMetLeuGly 182

Qy        648 ATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCG 707
Db        183 ValLeuAlaLeuProAlaValValLeuIleIleLeuValIlePheLeuProAsnSerPro 202

Qy        708 CGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCC 767
Db        203 ArgTrpLeuAlaGluLysGlyArgHisValGluAlaGluGluValLeuArgMetLeuArg 222

Qy        768 GACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCT 827
Db        223 AspThrSerGluLysAlaArgAspGluLeuAsnGluIleArgGluSerLeuLysLeu--- 241

Qy        828 GAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGG 887
Db        242 ---------------------------------LysGlnGlyGly-------------- 245

Qy        888 GTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGG 947
Db        246 ---Trp------AlaLeuPheLysValAsnArgAsnValArgArgAlaValPheLeuGly 262

Qy        948 ATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCT 1007
Db        263 MetLeuLeuGlnAlaMetGlnGlnPheThrGlyMetAsnIleIleMetTyrTyrAlaPro 282

Qy       1008 CTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCG 1067
Db        283 ArgIlePheLysMetAlaGlyPheThrThrThrGluGlnGlnMetValAlaThrLeuVal 302

Qy       1068 TTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGG 1127
Db        303 ValGlyLeuThrPheMetPheAlaThrPheIleAlaValPheThrValAspLysAlaGly 322

Qy       1128 CGGCGGCCGCTGTTG---------------CTGGGCAGCACGGGCGGGATAATCCTCTCC 1172
Db        323 ArgLysProAlaLeuLysIleGlyPheSerValMetAlaIleGlyThrLeuValLeuGly 342

Qy       1173 ------CTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAG 1226
Db        343 TyrCysLeuMetGlnPheAspAsnGlyThrAlaSerSerGly----------------- 356

Qy       1227 ATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCC 1286
Db        357 LeuSerTrp---------LeuSerValGlyMetThrMetMetCysIleAlaGlyTyrAla 373

Qy       1287 ATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGC 1346
Db        374 MetSerAlaAlaProValValTrpIleLeuCysSerGluIleGlnProLeuLysCysArg 393

Qy       1347 GCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATG 1406
Db        394 AspPheGlyIleThrCysSerThrThrThrAsnTrpValSerAsnMetIleIleGlyAla 413

Qy       1407 ACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGC 1466
Db        414 ThrPheLeuThrLeuLeuAspAlaIleGlyAlaAlaGlyThrPheTrpLeuTyrThrAla 433

Qy       1467 ATC---GCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGG 1523
Db        434 LeuAsnValAlaPheIleGlyValThrPheTrp---LeuIleProGluThrLysAsnVal 452

Qy       1524 ACGCTGGAGGAGATGAGCAAG 1544
Db        453 ThrLeuGluHisIleGluArg 459
```

Search completed: June 30, 2004, 18:53:07
Job time : 65 secs

GenCore version 5.1.6

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 30, 2004, 18:37:09 ; Search time 29.5 Seconds
(without alignments)
7120.373 Million cell updates/sec

Title: US-10-051-902A-21
Perfect score: 3614
Sequence: 1 cttacatgtaagctcgtgcc.....aaaaaaaaaaaaaaaaaaaa 2017

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 283362

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10051902/runat_30062004_164722_20400/app_query.fasta_1.2183
-DB=SwissProt_42 -QFMT=fastan -SUFFIX=rsp -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10051902_@CGN_1_1_22_@runat_30062004_164722_20400 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	654.5	18.1	629	1	MYCT_HUMAN	Q96qe2 homo sapien
2	609.5	16.9	461	1	CSBC_BACSU	P46333 bacillus su
3	601	16.6	557	1	ITR2_SCHPO	P87110 schizosacch
4	584.5	16.2	472	1	ARAE_KLEOX	P45598 klebsiella
5	582	16.1	457	1	XYLT_LACBR	O52733 lactobacill
6	579.5	16.0	472	1	ARAE_ECOLI	P09830 escherichia
7	576	15.9	584	1	ITR1_YEAST	P30605 saccharomyc
8	574	15.9	491	1	XYLE_ECOLI	P09098 escherichia
9	572.5	15.8	464	1	GALP_ECOLI	P37021 escherichia
10	572.5	15.8	612	1	ITR2_YEAST	P30606 saccharomyc
11	555.5	15.4	482	1	YFIG_BACSU	P54723 bacillus su
12	543.5	15.0	522	1	STP1_ARATH	P23586 arabidopsis
13	539	14.9	534	1	HUP3_CHLKE	Q39525 chlorella k
14	538.5	14.9	534	1	HUP1_CHLKE	P15686 chlorella k
15	531.5	14.7	468	1	GLCP_SYNY3	P15729 synechocyst
16	525.5	14.5	464	1	ARAE_BACSU	P96710 bacillus su
17	525	14.5	575	1	ITR1_SCHPO	Q10286 schizosacch
18	523.5	14.5	547	1	GTR1_LEIDO	Q01440 leishmania
19	505.5	14.0	510	1	HEX6_RICCO	Q07423 ricinus com
20	503.5	13.9	523	1	STC_RICCO	Q41144 ricinus com
21	502	13.9	522	1	STA_RICCO	Q10710 ricinus com
22	486	13.4	592	1	HXT5_YEAST	P38695 saccharomyc
23	484	13.4	541	1	GT10_HUMAN	O95528 homo sapien
24	483.5	13.4	541	1	HXT2_YEAST	P23585 saccharomyc
25	482.5	13.4	490	1	GTR1_CHICK	P46896 gallus gall
26	475	13.1	477	1	GTR8_MOUSE	Q9jif3 mus musculu
27	468.5	13.0	570	1	HXT1_YEAST	P32465 saccharomyc
28	467.5	12.9	537	1	QAY_NEUCR	P11636 neurospora
29	465.5	12.9	496	1	GTR3_CHICK	P28568 gallus gall
30	463	12.8	540	1	HUP2_CHLKE	Q39524 chlorella k
31	460.5	12.7	473	1	GLF_ZYMMO	P21906 zymomonas m
32	460.5	12.7	492	1	GTR1_RABIT	P13355 oryctolagus
33	456.5	12.6	492	1	GTR1_RAT	P11167 rattus norv
34	455.5	12.6	478	1	GTR8_RAT	Q9jjz1 rattus norv
35	455.5	12.6	492	1	GTR1_MOUSE	P17809 mus musculu
36	455.5	12.6	569	1	HXT8_YEAST	P40886 saccharomyc
37	454.5	12.6	492	1	GTR1_HUMAN	P11166 homo sapien
38	454.5	12.6	567	1	RAG1_KLULA	P18631 kluyveromyc
39	453	12.5	477	1	GTR8_HUMAN	Q9ny64 homo sapien
40	452.5	12.5	492	1	GTR1_BOVIN	P27674 bos taurus
41	451	12.5	546	1	HXT0_YEAST	P43581 saccharomyc
42	451	12.5	566	1	KHT2_KLULA	P53387 kluyveromyc
43	448	12.4	509	1	GTR4_HUMAN	P14672 homo sapien
44	445.5	12.3	495	1	GTR3_CANFA	P47842 canis famil
45	445.5	12.3	567	1	HXT9_YEAST	P40885 saccharomyc

ALIGNMENTS

RESULT 1
MYCT_HUMAN
ID MYCT_HUMAN STANDARD; PRT; 629 AA.
AC Q96QE2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Proton myo-inositol co-transporter (Hmit).
GN SLC2A13.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21391710; PubMed=11500374;
RA Uldry M., Ibberson M., Horisberger J.-D., Chatton J.-Y.,
RA Riederer B.M., Thorens B.;
RT "Identification of a mammalian H(+)-myo-inositol symporter expressed
RT predominantly in the brain.";
RL EMBO J. 20:4467-4477(2001).
CC -!- FUNCTION: H(+)-myo-inositol co-transporter. Can also transport
CC related stereoisomers.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Predominantly expressed in the brain.
CC -!- PTM: Glycosylated.
CC -!- SIMILARITY: Belongs to the sugar transporter family.
CC ----------

CC ----------
DR EMBL; AJ315644; CAC51116.1; -.
DR Genew; HGNC:15956; SLC2A13.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR005829; Sug_transporter.
DR InterPro; IPR003663; Sugar_transpt.

```
DR   Pfam; PF00083; sugar_tr; 1.
DR   PRINTS; PR00171; SUGRTRNSPORT.
DR   TIGRFAMs; TIGR00879; SP; 1.
DR   PROSITE; PS50850; MFS; 1.
DR   PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR   PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW   Transmembrane; Sugar transport; Glycoprotein.
FT   DOMAIN        1     57       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     58     78       1 (POTENTIAL).
FT   DOMAIN       79    106       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    107    127       2 (POTENTIAL).
FT   DOMAIN      128    129       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    130    150       3 (POTENTIAL).
FT   DOMAIN      151    159       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    160    180       4 (POTENTIAL).
FT   DOMAIN      181    193       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    194    214       5 (POTENTIAL).
FT   DOMAIN      215    220       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    221    241       6 (POTENTIAL).
FT   DOMAIN      242    305       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    306    326       7 (POTENTIAL).
FT   DOMAIN      327    344       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    345    365       8 (POTENTIAL).
FT   DOMAIN      366    374       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    375    395       9 (POTENTIAL).
FT   DOMAIN      396    489       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    490    510       10 (POTENTIAL).
FT   DOMAIN      511    530       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    531    551       11 (POTENTIAL).
FT   DOMAIN      552    554       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    555    575       12 (POTENTIAL).
FT   DOMAIN      576    629       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD    414    414       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    439    439       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    466    466       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   629 AA;  68203 MW;  8AD79C56C790814E CRC64;
```

Alignment Scores:

Pred. No.:	1.8e-24	Length:	629
Score:	654.50	Matches:	177
Percent Similarity:	44.01%	Conservative:	95
Best Local Similarity:	28.64%	Mismatches:	213
Query Match:	18.11%	Indels:	133
DB:	1	Gaps:	13

US-10-051-902A-21 (1-2017) x MYCT_HUMAN (1-629)

```
Qy      12 GCTCGTGCCGGCACGAGCTTACACTCGACCGCCACTACTGTACACGGCCCA--------- 62
           |||   :::   |||||||||   |||                    |||
Db      26 AlaGluSerSerThrSerLeuGlnSerAlaGlyAlaGlyGlyGlyGlyValGlyAspLeu 45

Qy      63 GAGCGAGCCTCCTCCTCCTCTGCACCACCGGAGATGGCTTCCGCCGCGCTGCCGGAGGCC 122
           |||||||||:::                :::       |||
Db      46 GluArgAlaAlaArgArgGlnPheGlnGlnAspGluThrProAla-------------- 60

Qy     123 GTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCC 182
                                       |||   :::     |||:::   ::::::
Db      61 ------------------------------PheValTyrValValAlaValPheSerAla 70

Qy     183 ATGACCTCCATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACATC 242
           :::         |||   |||||||||   ||||||:::|||||||||   |||   :::
Db      71 LeuGlyGlyPheLeuPheGlyTyrAspThrGlyValValSerGlyAlaMetLeuLeuLeu 90

Qy     243 AAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTC 302
           |||:::     ::::::              |||:::|||:::
Db      91 LysArgGlnLeuSerLeuAspAlaLeuTrpGlnGluLeuLeuValSerSerThrValGly 110

Qy     303 TACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTAC 362
              :::   :::   :::   |||   |||      :::         |||||||||
Db     111 AlaAlaAlaValSerAlaLeuAlaGlyGlyAlaLeuAsnGlyValPheGlyArgArgAla 130

Qy     363 ACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTC 422
           |||:::   |||:::   :::|||   ||||||      ::::::      |||
```

```
Db     131 AlaIleLeuLeuAlaSerAlaLeuPheThrAlaGlySerAlaValLeuAlaAlaAlaAsn 150

Qy     423 AACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTC 482
           |||         |||:::   ||||||  |||   |||:::|||:::|||   |||
Db     151 AsnLysGluThrLeuLeuAlaGlyArgLeuValValGlyLeuGlyIleGlyIleAlaSer 170

Qy     483 ATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACG 542
           |||       |||||||||   |||||||||||||||   :::   ||||||   |||
Db     171 MetThrValProValTyrIleAlaGluValSerProProAsnLeuArgGlyArgLeuVal 190

Qy     543 TCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCT 602
           :::         :::||||||     |||                 |||   :::   |||
Db     191 ThrIleAsnThrLeuPheIleThrGlyGlyGlnPhePheAlaSerValValAspGlyAla 210

Qy     603 TTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCG 662
           ||||||   |||      :::   |||||||||   |||||||||:::         |||
Db     211 PheSerTyrLeuGlnLysAsp---GlyTrpArgTyrMetLeuGlyLeuAla***ValPro 229

Qy     663 TCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATG 722
           :::|||:::                  |||   :::|||||||||||||||||||:::
Db     230 AlaValIleGlnPhePheGlyPheLeuPheLeuProGluSerProArgTrpLeuIleGln 249

Qy     723 AAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAG 782
           ||||||:::                                              ::::::
Db     250 LysGlyGln----------------------------------------------ThrGlnLys 255

Qy     783 GCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGC 842
           |||         |||:::   ::::::            |||   ||||||   |||
Db     256 AlaArgArgIleLeuSerGlnMetArgGlyAsnGlnThrIleAspGluGluTyrAsp--- 274

Qy     843 GACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTG----------- 890
                   ::::::                  ||||||   |||
Db     275 ---------SerIleLysAsnAsnIleGluGluGluGluLysGluValGlySerAlaGly 291

Qy     891 ---------TGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTG 941
                      :::   |||         ||||||         ||||||   |||:::
Db     292 ProValIleCysArgMetLeuSerTyrProProThr---------ArgArgAlaLeuIle 308

Qy     942 TCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTAC 1001
               |||   |||:::     ||||||          ||||||:::::::::::::|||
Db     309 ValGlyCysGlyLeuGlnMetPheGlnGlnLeuSerGlyIleAsnThrIleMetTyrTyr 328

Qy    1002 AGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGCACCACT 1061
           |||       :::   :::       |||:::   :::|||:::       :::
Db     329 SerAlaThrIleLeuGlnMetSerGlyValGluAspAspArgLeuAlaIleTrpLeuAla 348

Qy    1062 TGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGC 1121
                         |||       :::|||   |||:::       :::   ::::::
Db     349 SerValThrAlaPheThrAsnPheIlePheThrLeuValGlyValTrpLeuValGluLys 368

Qy    1122 GTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGC 1181
           ||||||||||||   |||       ||||||       |||       ::::::||||||
Db     369 ValGlyArgArgLysLeuThrPheGlySerLeuAlaGlyThrThrValAlaLeuIleIle 388

Qy    1182 CTCGGCGCCGGGCTCACCGTCGTCGGCCAG------------------------------ 1211
           |||       |||       :::       |||
Db     389 LeuAlaLeuGlyPheValLeuSerAlaGlnValSerProArgIleThrPheLysProIle 408

Qy    1211 ---------------------------------------------------------- 1211

Db     409 AlaProSerGlyGlnAsnAlaThrCysThrArgTyrSerTyrCysAsnGluCysMetLeu 428

Qy    1212 CACCCCGACGCC---------------------------------------------- 1223
              ||||||
Db     429 AspProAspCysGlyPheCys***LysMetAsnLysSerThrValIleAspSerSerCys 448

Qy    1223 ---------------------------------------------------------- 1223

Db     449 ValProValAsnLysAlaSerThrAsnGluAlaAlaTrpGlyArgCysGluAsnGluThr 468

Qy    1224 --------------AAGATACCTTGGGCCATCGGCCTA---------------------- 1247
                           |||   ||||||
Db     469 LysPheLysThrGluAspIlePheTrpAlaTyrAsnPheCysProThrProTyrSerTrp 488
```

```
Qy       1248 ---AGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATC 1304
                 ::::::         :::   |||:::   ||||||:::   |||:::||||||:::
Db        489 ThrAlaLeuLeuGlyLeuIleLeuTyrLeuValPhePheAlaProGlyMetGlyProMet 508

Qy       1305 ACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGC------TGC 1358
                |||      :::|||||||||:::||||||      |||:::   |||      |||
Db        509 ProTrpThrValAsnSerGluIleTyrProLeuTrpAlaArgSerThrGlyAsnAlaCys 528

Qy       1359 TCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCG 1418
                |||   |||:::      |||   :::   :::   ::::::|||:::|||||||||
Db        529 SerSerGlyIle------AsnTrpIlePheAsnValLeuValSerLeuThrPheLeuHis 546

Qy       1419 CTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTC 1478
                ::::::   :::|||      |||:::||||||||||||:::|||   ||||||:::
Db        547 ThrAlaGluTyrLeuThrTyrTyrGlyAlaPhePheLeuTyrAlaGlyPheAlaAlaVal 566

Qy       1479 GCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATG 1538
                 :::|||   |||      ||||||||||||:::|||:::   |||||||||:::
Db        567 GlyLeuLeuPheIleTyrGlyCysLeuProGluThrLysGlyLysLysLeuGluGluIle 586

Qy       1539 AGCAAGCTGTTCGGCGAC---------ACGGCCGCCGCCTCGGAATCAGACGAG 1583
                 ||||||   :::         |||         |||:::|||||||||
Db        587 GluSerLeuPheAspAsnArgLeuCysThrCysGlyThrSerAspSerAspGlu 604
```

RESULT 2
CSBC_BACSU
ID CSBC_BACSU STANDARD; PRT; 461 AA.
AC P46333; O32289;
DT 01-NOV-1995 (Rel. 32, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Probable metabolite transport protein csbC.
GN CSBC OR SS92BR OR BSU39810.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / BGSC1A1;
RX MEDLINE=96093926; PubMed=7584049;
RA Yoshida K.-I., Seki S., Fujimura M., Miwa Y., Fujita Y.;
RT "Cloning and sequencing of a 36-kb region of the Bacillus subtilis
RT genome between the gnt and iol operons.";
RL DNA Res. 2:61-69(1995).
RN [2]
RP REVISIONS.
RA Fujita Y., Shibayama T., Ishio I., Aoyama D., Yoshida K.-I.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis.";
RL Nature 390:249-256(1997).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=99303315; PubMed=10376822;
RA Akbar S., Lee S.Y., Boylan S.A., Price C.W.;
RT "Two genes from Bacillus subtilis under the sole control of the
RT general stress transcription factor sigmaB.";
RL Microbiology 145:1069-1078(1999).
CC -!- FUNCTION: Could serve either a nutritional or an osmotic
CC protection function.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!- SIMILARITY: Belongs to the sugar transporter family.
CC ----------
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ----------
DR EMBL; AB005554; BAA21604.1; -.
DR EMBL; Z99124; CAB16017.1; -.
DR PIR; D70073; D70073.
DR SubtiList; BG11360; csbC.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR005829; Sug_transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR Pfam; PF00083; sugar_tr; 1.
DR PRINTS; PR00171; SUGRTRNSPORT.
DR TIGRFAMs; TIGR00879; SP; 1.
DR PROSITE; PS50850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Transport; Transmembrane; Complete proteome.

FT	DOMAIN	1	14	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	15	35	1 (POTENTIAL).
FT	DOMAIN	36	38	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	39	59	2 (POTENTIAL).
FT	DOMAIN	60	76	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	77	97	3 (POTENTIAL).
FT	DOMAIN	98	104	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	105	125	4 (POTENTIAL).
FT	DOMAIN	126	139	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	140	160	5 (POTENTIAL).
FT	DOMAIN	161	163	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	164	184	6 (POTENTIAL).
FT	DOMAIN	185	241	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	242	262	7 (POTENTIAL).
FT	DOMAIN	263	280	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	281	301	8 (POTENTIAL).
FT	DOMAIN	302	308	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	309	329	9 (POTENTIAL).
FT	DOMAIN	330	341	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	342	362	10 (POTENTIAL).
FT	DOMAIN	363	378	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	379	399	11 (POTENTIAL).
FT	DOMAIN	400	402	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	403	423	12 (POTENTIAL).
FT	DOMAIN	424	461	CYTOPLASMIC (POTENTIAL).
FT	CONFLICT	400	401	SA -> RP (IN REF. 1).

SQ SEQUENCE 461 AA; 50140 MW; AD59CB5DD6372FE0 CRC64;

```
Alignment Scores:
Pred. No.:                2.5e-22        Length:          461
Score:                    609.50         Matches:         149
Percent Similarity:       49.59%         Conservative:    93
Best Local Similarity:    30.53%         Mismatches:      195
Query Match:              16.86%         Indels:          51
DB:                       1              Gaps:            11

US-10-051-902A-21 (1-2017) x CSBC_BACSU (1-461)

Qy        180 TCCATGACCTCCATCCTCCTCGGCTACGATATCGGGGGTGATGAGCGGGGCGTCGCTGTAC 239
              ::::::      :::||| ||||||||  ||||||:::|||||||||   |||:::
Db         15 AlaLeuGlyGlyLeuLeuTyrGlyTyrAspThrGlyValIleSerGlyAlaLeuLeuPhe 34

Qy        240 ATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAAC 299
              |||      |||      ::::::              ::::::::   :::|||
Db         35 IleAsnAsnAspIleProLeuThrThrLeuThrGluGlyLeuValValSerMetLeuLeu 54

Qy        300 CTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGG 359
              |||.  ::::::  ||||||      :::|||      ||||||      ||||||||
Db         55 LeuGlyAlaIlePheGlySerAlaLeuSerGlyThrCysSerAspArgTrpGlyArgArg 74

Qy        360 TACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGGSGTTCCTCATGGGGTTCGCC 419
                 :::       ::::::||||||     |||                 |||:::
Db         75 LysValValPheValLeuSerIleIlePheIleIleGlyAlaLeuAlaCysAlaPheSer 94

Qy        420 GTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCG 479
                    ||||||:::      |||   :::   |||:::   ||||||   :::
Db         95 GlnThrIleGlyMetLeuIleAlaSerArgValIleLeuGlyLeuAlaValGlyGlySer 114

Qy        480 CTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTG 539
               :::   |||||||||   :::|||::::::|||          ||||||   |||
Db        115 ThrAlaLeuValProValTyrLeuSerGluMetAlaProThrLysIleArgGlyThrLeu 134

Qy        540 ACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTAT 599
              :::       :::   |||      ||||||||||   |||:::   ||||||
Db        135 GlyThrMetAsnAsnLeuMetIleValThrGlyIleLeuLeuAlaTyrIleValAsnTyr 154

Qy        600 GCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCG 659
              |||:::      |||          ||||||   |||:::|||:::   |||
Db        155 LeuPheThr------ProPheGlu---AlaTrpArgTrpMetValGlyLeuAlaAlaVal 171

Qy        660 CCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTC 719
              |||:::|||||||||   :::   :::      ||||||||||||||||||||||||||
Db        172 ProAlaValLeuLeuLeuIleGlyIleAlaPheMetProGluSerProArgTrpLeuVal 191

Qy        720 ATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAG 779
              :::|||         :::|||:::   ::::::      |||   |||
Db        192 LysArgGlySerGluGluGluAlaArgArgIleMetAsnIleThrHisAsp--------- 208

Qy        780 GAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGAC 839
                                                          |||:::::::::::
Db        209 ----------------------------------------------ProLysAspIleGlu 213

Qy        840 GGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAG 899
              ::::::      :::   |||:::|||          :::                |||||
Db        214 MetGluLeuAlaGluMet---LysGlnGlyGluAlaGluLys-----------LysGlu 228

Qy        900 CTCATCCTGTCCCCGACCCCGGCC------ATGCGGCGCATCCTGCTGTCCGGGATCGGC 953
              |||              |||      :::|||   :::||||||   |||:::|||
Db        229 ThrThrLeuGlyValLeuLysAlaLysTrpIleArgProMetLeuLeuIleGlyValGly 248

Qy        954 ATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTG 1013
              :::      ||||||   |||:::||||||::::::|||::::::|||:::|||   :::
Db        249 LeuAlaIlePheGlnGlnAlaValGlyIleAsnThrValIleTyrTyrAlaProThrIle 268

Qy       1014 TTCAAGAGCCCCCGGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGT 1073
              |||      ||||||                  ||||||   |||      |||
Db        269 PheThrLysAlaGlyLeuGlyThrSerAlaSerAlaLeuGly---ThrMetGlyIleGly 287

Qy       1074 GTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGG 1133
              :::      :::      :::   |||        ||||||   |||||||||:::

Db        288 IleLeuAsnValIleMetCysIleThrAlaMetIleLeuIleAspArgValGlyArgLys 307

Qy       1134 CCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGG 1193
                 ||||||:::      :::   ||||||  ||||||||      |||
Db        308 LysLeuLeuIleTrpGlySerValGlyIleThrLeuSerLeuAlaAlaLeuSerGlyVal 327

Qy       1194 CTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATC 1253
              |||      :::|||             |||         |||         :::::::::
Db        328 LeuLeuThrLeuGlyLeu------SerAlaSerThrAlaTrp---------MetThrVal 342

Qy       1254 GCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTG 1313
                          |||:::   |||:::            ||||||:::   ||||||
Db        343 ValPheLeuGlyValTyrIleValPheTyrGlnAlaThrTrpGlyProValValTrpVal 362

Qy       1314 TACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCC 1373
                       |||:::||||||   :::   |||
Db        363 LeuMetProGluLeuPheProSerLysAlaArgGlyAlaAlaThrGlyPheThrThrLeu 382

Qy       1374 GCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATC 1433
                          :::   ::::::|||:::   |||      :::      |||:::
Db        383 ValLeuSerAlaAlaAsnLeuIleValSerLeuValPheProLeuMetLeuSerAlaMet 402

Qy       1434 ACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTC 1493
                 |||         |||   ::::::|||   |||      |||::::::   |||
Db        403 GlyIleAlaTrpValPheMetValPheSerValIleCysLeuLeuSerPhePhePheAla 422

Qy       1494 TACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAG------------ATGAGC 1541
              :::      :::|||||||||:::|||::::::|||||||||            :::
Db        423 PheTyrMetValProGluThrLysGlyLysSerLeuGluGluIleGluAlaSerLeuLys 442

Qy       1542 AAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAGCCAAGGAGAAGAAG 1601
                 |||   |||                                        |||:::||||||
Db        443 LysArgPhe----------------------------------------LysLysLysLys 449

Qy       1602 AAGGTGGAAATGGCCGCCACTAAC 1625
                       :::          |||
Db        450 SerThrGlnAsnGlnValLeuAsn 457

RESULT 3
ITR2_SCHPO
ID   ITR2_SCHPO     STANDARD;      PRT;   557 AA.
AC   P87110; P78901;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Myo-inositol transporter 2.
GN   ITR2 OR SPAC20G8.03.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=968 h90;
RX   MEDLINE=98228265; PubMed=9560432;
RA   Niederberger C., Graub R., Schweingruber A.-M., Fankhauser H.,
RA   Rusu M., Poitelea M., Edenharter L., Schweingruber M.E.;
RT   "Exogenous inositol and genes responsible for inositol transport are
RT   required for mating and sporulation in Shizosaccharomyces pombe.";
RL   Curr. Genet. 33:255-261(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RX   MEDLINE=21848401; PubMed=11859360;
RA   Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA   Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA   Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA   Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA   Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA   Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA   James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA   Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
```

```
RA   Oliver K., O'Neil S., Pearson D., Quail M.A., Rabbinowitsch E.,
RA   Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA   Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA   Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA   Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA   Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA   Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA   Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA   Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA   Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA   Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA   Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA   Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA   Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA   Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA   Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT   "The genome sequence of Schizosaccharomyces pombe.";
RL   Nature 415:871-880(2002).
RN   [3]
RP   SEQUENCE OF 166-541 FROM N.A.
RC   STRAIN=PR745;
RX   MEDLINE=98162722; PubMed=9501991;
RA   Yoshioka S., Kato K., Nakai K., Okayama H., Nojima H.;
RT   "Identification of open reading frames in Schizosaccharomyces pombe
RT   cDNAs.";
RL   DNA Res. 4:363-369(1997).
CC   -!- FUNCTION: TRANSPORTER FOR MYO-INOSITOL.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to the sugar transporter family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X99105; -; NOT_ANNOTATED_CDS.
DR   EMBL; Z95334; CAB08597.1; -.
DR   EMBL; D89252; BAA13913.1; -.
DR   PIR; T38125; T38125.
DR   GeneDB_SPombe; SPAC20G8.03; -.
DR   InterPro; IPR007114; MFS.
DR   InterPro; IPR005828; Sub_transporter.
DR   InterPro; IPR005829; Sug_transporter.
DR   InterPro; IPR003663; Sugar_transpt.
DR   Pfam; PF00083; sugar_tr; 1.
DR   PRINTS; PR00171; SUGRTRNSPORT.
DR   TIGRFAMs; TIGR00879; SP; 1.
DR   PROSITE; PS50850; MFS; 1.
DR   PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.
DR   PROSITE; PS00217; SUGAR_TRANSPORT_2; FALSE_NEG.
KW   Transmembrane; Sugar transport.
FT   DOMAIN        1     76       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     77     97       1 (POTENTIAL).
FT   DOMAIN       98     99       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    100    120       2 (POTENTIAL).
FT   DOMAIN      121    123       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    124    144       3 (POTENTIAL).
FT   DOMAIN      145    157       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    158    178       4 (POTENTIAL).
FT   DOMAIN      179    180       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    181    201       5 (POTENTIAL).
FT   DOMAIN      202    209       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    210    230       6 (POTENTIAL).
FT   DOMAIN      231    240       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    241    261       7 (POTENTIAL).
FT   DOMAIN      262    367       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    368    388       8 (POTENTIAL).
FT   DOMAIN      389    396       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    397    417       9 (POTENTIAL).
FT   DOMAIN      418    432       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    433    453       10 (POTENTIAL).
FT   DOMAIN      454    468       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    469    489       11 (POTENTIAL).
FT   DOMAIN      490    498       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    499    519       12 (POTENTIAL).
FT   DOMAIN      520    557       CYTOPLASMIC (POTENTIAL).
FT   VARIANT     271    271       N -> H (IN STRAIN PR745).
FT   VARIANT     428    428       N -> H (IN STRAIN PR745).
FT   VARIANT     436    436       L -> I (IN STRAIN PR745).
FT   VARIANT     526    526       S -> F (IN STRAIN PR745).
SQ   SEQUENCE   557 AA;  61136 MW;  20875EC11B153175 CRC64;

Alignment Scores:
Pred. No.:                6.42e-22        Length:          557
Score:                    601.00          Matches:         160
Percent Similarity:       51.22%          Conservative:    112
Best Local Similarity:    30.13%          Mismatches:      211
Query Match:              16.63%          Indels:          48
DB:                       1               Gaps:            16

US-10-051-902A-21 (1-2017) x ITR2_SCHPO (1-557)

Qy          93 GAGATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGG 152
               |||   |||:::|||      |||:::|||           :::
Db          52 GluGluAlaAsnAlaThrAspProGlnAlaAsnGluValAlaAspGluAsnGlyGluGly 71

Qy         153 TTC----------------------GCCTTCGCCTGCGCCATCCTCGCCTCCATGACC 188
               |||                        :::      :::   :::|||   ::::::
Db          72 PheGluAlaGluLysIleSerSerTrpIleTrpValLeuSerAlaValAlaGlyIleSer 91

Qy         189 TCCATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAG 248
                  :::|||   |||||||||   ||||||:::|||||||||          :::
Db          92 GlyLeuLeuPheGlyTyrAspThrGlyValIleSerGlyAlaLeuAlaValLeuGlySer 111

Qy         249 GACTTCAAC------ATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTC 302
               |||          :::|||   |||:::   |||::::::          :::
Db         112 AspLeuGlyHisValLeuSerSerGlyGlnLysGluLeuIleThrSerAlaThrSerPhe 131

Qy         303 TACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTAC 362
                  :::||||||   :::      :::|||      :::||||||:::||||||:::
Db         132 AlaAlaLeuIleSerAlaThrThrSerGlyTrpLeuAlaAspTrpValGlyArgLysArg 151

Qy         363 ACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTC 422
                  ::::::   |||      ||||||     |||      :::|||      :::
Db         152 LeuLeuLeuCysAlaAspAlaIlePheValIleGlySerValIleMetAlaAlaSerArg 171

Qy         423 AACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTC 482
               |||   ||||||::::::   ||||||||||:::   |||   |||:::|||
Db         172 AsnValAlaMetMetValValGlyArgPheIleValGlyTyrGlyIleGlyLeuThrSer 191

Qy         483 ATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACG 542
               :::|||   |||:::|||      |||::::::||||||      ||||||   |||
Db         192 LeuIleValProMetTyrIleThrGluLeuAlaProAlaArgLeuArgGlyArgLeuVal 211

Qy         543 TCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCT 602
                        |||||||||      |||   |||:::   |||      |||   |||
Db         212 IleIleTyrValValPheIleThrGlyGlyGlnLeuIleAlaTyrSerLeuAsnAlaAla 231

Qy         603 TTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCG 662
               |||          ::::::   ||||||||||||||   ||||||||||||||||
Db         232 PheGluHis------ValHisGlnGlyTrpArgIleMetPheGlyIleGlyAlaAlaPro 249

Qy         663 TCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATG 722
               ::::::       :::   :::          |||||||||||||:::|||:::
Db         250 AlaLeuGlyGlnLeuIleSerLeuPheTrpThrProGluSerProArgTyrLeuLeuArg 269

Qy         723 AAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAG--- 779
                     :::                 :::|||   :::            |||:::
Db         270 HisAsnHisValGluLysValTyrLysIleLeuSerArgIleHisProGluAlaLysPro 289

Qy         780 ---GAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTC 836
                  |||   |||   :::::::::   |||:::
Db         290 AlaGluIleAlaTyrLysValSerLeuIleGln---------------------------- 300
```

```
Qy        837 GACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAG---AAGCGGGTGTGG 893
Db        301 GluGlyValLysValAspProGlu---------GlyAsnLysPheGlnHisPhePhe 317

Qy        894 AAGGAGCTC---ATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATC 950
Db        318 HisSerLeuLysValLeuPheThrValProSerAsnArgArgSerLeuPheIleGlyCys 337

Qy        951 GGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTC 1010
Db        338 PheLeuGlnTrpPheGlnGlnPheSerGlyThrAsnAlaIleGlnTyrPheSerAlaIle 357

Qy       1011 GTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTC 1070
Db        358 IlePheGlnSerValGlyPheLysAsnSer---------IleSerValSerIleValVal 374

Qy       1071 GGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGG 1130
Db        375 GlyAlaThrAsnPheValPheThrIleValAlaPheMetPheIleAspArgIleGlyArg 394

Qy       1131 CGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCC 1190
Db        395 ArgArgIleLeuLeu---CysThrSerAlaValMetIleAla-----------GlyLeu 409

Qy       1191 GGGCTCACCGTCGTCGGCCAGCAC--------------------CCCGACGCCAAGATA 1229
Db        410 AlaLeuCysAlaIleAlaTyrHisPheLeuProAlaAspThrThrGlnAsnThrAsnSer 429

Qy       1230 CCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATC 1289
Db        430 GlyTrpGlnTyr---ValValLeuAlaSerIleIleIlePheLeuAlaSerTyrAlaSer 448

Qy       1290 GGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCG 1349
Db        449 GlyIleGlyAsnIleProTrp---GlnGlnAlaGluLeuPheProMetGluValArgAla 467

Qy       1350 CTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACC 1409
Db        468 LeuGlyAlaGlyPheSerThrAlaIleAsnTrpValGlyAsnLeuIleIleSerAlaSer 487

Qy       1410 TTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATC 1469
Db        488 PheLeuThrMetMetGluSerIleThrProThrGlyThrPheAlaLeuPheAlaGlyPhe 507

Qy       1470 GCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTG 1529
Db        508 CysPheValGlyLeuValThrSerTyrPheThrTyrProGluLeuAlaGlyMetSerIle 527

Qy       1530 GAGGAGATGAGCAAGCTGTTC------GGCGACACGGCCGCCGCCTCGGAATCAGACGAG 1583
Db        528 GluAsnIleHisLysLeuLeuGluLysGlyPheTrpGlnAlaValLysGluSerThrLys 547

Qy       1584 CCAGCCAAGGAGAAGAAGAAGGTGGAAATGGCC 1616
Db        548 ArgValArg---LysGlyArgIleAspGluAla 557
```

```
RESULT 4
ARAE_KLEOX
ID   ARAE_KLEOX     STANDARD;      PRT;   472 AA.
AC   P45598;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Arabinose-proton symporter (Arabinose transporter).
GN   ARAE.
OS   Klebsiella oxytoca.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Klebsiella.
OX   NCBI_TaxID=571;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=8017;
RX   MEDLINE=95394866; PubMed=7665532;
RA   Shatwell K.P., Charalambous B.M., McDonald T.P., Henderson P.J.F.;
RT   "Cloning, sequencing, and expression of the araE gene of Klebsiella
RT   oxytoca 8017, which encodes arabinose-H+ symport activity.";
RL   J. Bacteriol. 177:5379-5380(1995).
CC   -!- FUNCTION: Uptake of arabinose across the boundary membrane with
CC       the concomitant export of a proton (symport system).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC   -!- SIMILARITY: Belongs to the sugar transporter family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X79598; CAA56110.1; -.
DR   PIR; S47089; S47089.
DR   InterPro; IPR007114; MFS.
DR   InterPro; IPR005828; Sub_transporter.
DR   InterPro; IPR005829; Sug_transporter.
DR   InterPro; IPR003663; Sugar_transpt.
DR   Pfam; PF00083; sugar_tr; 1.
DR   PRINTS; PR00171; SUGRTRNSPORT.
DR   TIGRFAMs; TIGR00879; SP; 1.
DR   PROSITE; PS50850; MFS; 1.
DR   PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR   PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW   Transport; Sugar transport; Transmembrane; Inner membrane; Symport.
FT   DOMAIN        1     29       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     30     50       1 (POTENTIAL).
FT   DOMAIN       51     63       PERIPLASMIC (POTENTIAL).
FT   TRANSMEM     64     84       2 (POTENTIAL).
FT   DOMAIN       85     91       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     92    112       3 (POTENTIAL).
FT   DOMAIN      113    114       PERIPLASMIC (POTENTIAL).
FT   TRANSMEM    115    135       4 (POTENTIAL).
FT   DOMAIN      136    154       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    155    175       5 (POTENTIAL).
FT   DOMAIN      176    178       PERIPLASMIC (POTENTIAL).
FT   TRANSMEM    179    199       6 (POTENTIAL).
FT   DOMAIN      200    257       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    258    278       7 (POTENTIAL).
FT   DOMAIN      279    297       PERIPLASMIC (POTENTIAL).
FT   TRANSMEM    298    318       8 (POTENTIAL).
FT   DOMAIN      319    325       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    326    346       9 (POTENTIAL).
FT   DOMAIN      347    361       PERIPLASMIC (POTENTIAL).
FT   TRANSMEM    362    382       10 (POTENTIAL).
FT   DOMAIN      383    404       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    405    425       11 (POTENTIAL).
FT   DOMAIN      426    427       PERIPLASMIC (POTENTIAL).
FT   TRANSMEM    428    446       12 (POTENTIAL).
FT   DOMAIN      447    472       CYTOPLASMIC (POTENTIAL).
SQ   SEQUENCE   472 AA;  51732 MW;  410021E1BEE3D96E CRC64;
```

```
Alignment Scores:
Pred. No.:                3.91e-21        Length:          472
Score:                    584.50          Matches:         143
Percent Similarity:       52.25%          Conservative:    101
Best Local Similarity:    30.62%          Mismatches:      182
Query Match:              16.17%          Indels:          41
DB:                       1               Gaps:            10

US-10-051-902A-21 (1-2017) x ARAE_KLEOX (1-472)

Qy         168 GCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGGGGTGATGAGCGGG 227
Db          26 SerIleAlaAlaAlaValAlaGlyLeuLeuPheGlyLeuAspIleGlyValIleAlaGly 45

Qy         228 GCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATG 287
```

```
Db        46 AlaLeuProPheIleThrAspHisPheValLeuSerSerArgLeuGlnGluTrpValVal 65

Qy       288 GGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGG 347
Db        66 SerSerMetMetLeuGlyAlaAlaIleGlyAlaLeuPheAsnGlyTrpLeuSerPheArg 85

Qy       348 ATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGSGTTCCTC 407
Db        86 LeuGlyArgLysTyrSerLeuMetValGlyAlaValLeuPheValAlaGlySerValGly 105

Qy       408 ATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGC 467
Db       106 SerAlaPheAlaThrSerValGluMetLeuLeuValAlaArgIleValLeuGlyValAla 125

Qy       468 GTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCG 527
Db       126 ValGlyIleAlaSerTyrThrAlaProLeuTyrLeuSerGluMetAlaSerGluAsnVal 145

Qy       528 CGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTAC 587
Db       146 ArgGlyLysMetIleSerMetTyrGlnLeuMetValThrLeuGlyIleValMetAlaPhe 165

Qy       588 GTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGC 647
Db       166 LeuSerAspThrAlaPheSer---------TyrSerGlyAsnTrpArgAlaMetLeuGly 182

Qy       648 ATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCG 707
Db       183 ValLeuAlaLeuProAlaValValLeuIleIleLeuValIlePheLeuProAsnSerPro 202

Qy       708 CGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCC 767
Db       203 ArgTrpLeuAlaGluLysGlyArgHisValGluAlaGluGluValLeuArgMetLeuArg 222

Qy       768 GACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCT 827
Db       223 AspThrSerGluLysAlaArgAspGluLeuAsnGluIleArgGluSerLeuLysLeu--- 241

Qy       828 GAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGG 887
Db       242 ---------------------------------LysGlnGlyGly-------------- 245

Qy       888 GTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGG 947
Db       246 ---Trp------AlaLeuPheLysValAsnArgAsnValArgArgAlaValPheLeuGly 262

Qy       948 ATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCT 1007
Db       263 MetLeuLeuGlnAlaMetGlnGlnPheThrGlyMetAsnIleIleMetTyrTyrAlaPro 282

Qy      1008 CTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCG 1067
Db       283 ArgIlePheLysMetAlaGlyPheThrThrThrGluGlnGlnMetValAlaThrLeuVal 302

Qy      1068 TTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGG 1127
Db       303 ValGlyLeuThrPheMetPheAlaThrPheIleAlaValPheThrValAspLysAlaGly 322

Qy      1128 CGGCGGCCGCTGTTG---------------CTGGGCAGCACGGGCGGGATAATCCTCTCC 1172
Db       323 ArgLysProAlaLeuLysIleGlyPheSerValMetAlaIleGlyThrLeuValLeuGly 342

Qy      1173 ------CTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAG 1226
Db       343 TyrCysLeuMetGlnPheAspAsnGlyThrAlaSerSerGly----------------- 356

Qy      1227 ATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCC 1286
Db       357 LeuSerTrp---------LeuSerValGlyMetThrMetMetCysIleAlaGlyTyrAla 373

Qy      1287 ATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGC 1346
```

```
Db       374 MetSerAlaAlaProValValTrpIleLeuCysSerGluIleGlnProLeuLysCysArg 393

Qy      1347 GCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATG 1406
Db       394 AspPheGlyIleThrCysSerThrThrThrAsnTrpValSerAsnMetIleIleGlyAla 413

Qy      1407 ACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGC 1466
Db       414 ThrPheLeuThrLeuLeuAspAlaIleGlyAlaAlaGlyThrPheTrpLeuTyrThrAla 433

Qy      1467 ATC---GCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGG 1523
Db       434 LeuAsnValAlaPheIleGlyValThrPheTrp---LeuIleProGluThrLysAsnVal 452

Qy      1524 ACGCTGGAGGAGATGAGCAAG 1544
Db       453 ThrLeuGluHisIleGluArg 459
```

RESULT 5
XYLT_LACBR
ID XYLT_LACBR STANDARD; PRT; 457 AA.
AC O52733;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE D-xylose-proton symporter (D-xylose transporter).
GN XYLT.
OS Lactobacillus brevis.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1580;
RN [1]
RP SEQUENCE FROM N.A.
RA Chaillou S., Bor Y.-C., Batt C.A., Postma P., Pouwels P.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
CC -!- FUNCTION: Uptake of D-xylose across the boundary membrane with
CC the concomitant export of a proton (symport system).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to the sugar transporter family.
CC ----------
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ----------
DR EMBL; AF045552; AAC95127.1; -.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR005829; Sug_transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR Pfam; PF00083; sugar_tr; 1.
DR PRINTS; PR00171; SUGRTRNSPORT.
DR TIGRFAMs; TIGR00879; SP; 1.
DR PROSITE; PS50850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Transport; Sugar transport; Transmembrane; Symport.
FT DOMAIN 1 13 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 14 34 1 (POTENTIAL).
FT DOMAIN 35 45 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 46 66 2 (POTENTIAL).
FT DOMAIN 67 80 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 81 101 3 (POTENTIAL).
FT DOMAIN 102 103 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 104 124 4 (POTENTIAL).
FT DOMAIN 125 130 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 131 151 5 (POTENTIAL).
FT DOMAIN 152 163 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 164 184 6 (POTENTIAL).
FT DOMAIN 185 243 CYTOPLASMIC (POTENTIAL).

```
FT   TRANSMEM     244     264       7 (POTENTIAL).
FT   DOMAIN       265     280       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     281     301       8 (POTENTIAL).
FT   DOMAIN       302     309       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     310     330       9 (POTENTIAL).
FT   DOMAIN       331     337       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     338     358       10 (POTENTIAL).
FT   DOMAIN       359     379       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     380     400       11 (POTENTIAL).
FT   DOMAIN       401     401       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     402     422       12 (POTENTIAL).
FT   DOMAIN       423     457       CYTOPLASMIC (POTENTIAL).
SQ   SEQUENCE   457 AA;  49199 MW;  E097EB2B67A92F67 CRC64;

Alignment Scores:
Pred. No.:                5.13e-21      Length:          457
Score:                    582.00        Matches:         154
Percent Similarity:       49.80%        Conservative:    96
Best Local Similarity:    30.68%        Mismatches:      186
Query Match:              16.10%        Indels:          66
DB:                       1             Gaps:            13

US-10-051-902A-21 (1-2017) x XYLT_LACBR (1-457)

Qy        135 AAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATC 194
Db          2 ArgLysValSerThrGlyPheValTyr---------PhePheGlyAlaLeuGlyGlyLeu 18

Qy        195 CTCCTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTC 254
Db         19 LeuPheGlyTyrAspThrGlyValIleSerGlyAlaIleLeuPheIleGlnLysGlnMet 38

Qy        255 AACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATC 314
Db         39 AsnLeuGlySerTrpGlnGlnGlyTrpValValSerAlaValLeuLeuGlyAlaIleLeu 58

Qy        315 GGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTC 374
Db         59 GlyAlaAlaIleIleGlyProSerSerAspArgPheGlyArgArgLysLeuLeuLeuLeu 78

Qy        375 GCCGCCGTCATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATG 434
Db         79 SerAlaIleIlePhePheValGlyAlaLeuGlySerAlaPheSerProGluPheTrpThr 98

Qy        435 CTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCG 494
Db         99 LeuIleIleSerArgIleIleLeuGlyMetAlaValGlyAlaAlaSerAlaLeuIlePro 118

Qy        495 GTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAG 554
Db        119 ThrTyrLeuAlaGluLeuAlaProSerAspLysArgGlyThrValSerSerLeuPheGln 138

Qy        555 GTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTG 614
Db        139 LeuMetValMetThrGlyIleLeuLeuAlaTyrIleThrAsnTyrSerPheSer------ 156

Qy        615 CCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTC 674
Db        157 GlyPheTyrThrGlyTrpArgTrpMetLeuGlyPheAlaAlaIleProAlaAlaLeuLeu 176

Qy        675 GCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTC 734
Db        177 PheLeuGlyGlyLeuIleLeuProGluSerProArgPheLeuValLysSerGlyHisLeu 196

Qy        735 GCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCG------ 788
Db        197 AspGluAlaArgHisValLeuAsp---ThrMetAsnLysHisAspGlnValAlaValAsn 215

Qy        789 GAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTG 848
Db        216 LysGluIleAsnAspIleGlnGluSerAlaLysIleVal-------------------- 228

Qy        849 GTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTG 908

Db        229 ----------SerGlyGly----------TrpSerGluLeuPhe--- 236

Qy        909 TCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAG 968
Db        237 ------GlyLysMetValArgProSerLeuIleIleGlyIleGlyLeuAlaIlePheGln 254

Qy        969 CATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGA 1028
Db        255 GlnValMetGlyCysAsnThrValLeuTyrTyrAlaProThrIlePheThrAspValGly 274

Qy       1029 TTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTT 1088
Db        275 ----------PheGlyValSerAlaAlaLeu 281

Qy       1089 TTC----------ATCTTGTTGGCGACTTTCTTC 1112
Db        282 LeuAlaHisIleGlyIleGlyIlePheAsnValIleValThrAlaIleAlaValAlaIle 301

Qy       1113 ATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCC 1172
Db        302 MetAspLysIleAspArgLysLysIleValAsnIleGlyAlaValGlyMetGlyIleSer 321

Qy       1173 CTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCT 1232
Db        322 LeuPheValMetSerIleGlyMetLysPheSerGlyGlySerGlnThrAlaAlaIle--- 340

Qy       1233 TGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGC 1292
Db        341 ------------IleSerValIleAlaLeuThrValTyrIleAlaPhePheSerAlaThr 356

Qy       1293 CTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTG 1352
Db        357 TrpGlyProValMetTrpValMetIleGlyGluValPheProLeuAsnIleArgGlyLeu 376

Qy       1353 GGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTC 1412
Db        377 GlyAsnSerPheAlaSerValIleAsnTrpThrAlaAsnMetIleValSerLeuThrPhe 396

Qy       1413 CTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCC 1472
Db        397 ProSerLeuLeuAspPhePheGlyThr---GlySerLeuPheIleGlyTyrGlyIleLeu 415

Qy       1473 GCGCTCGCCTGGGTGTTCTTC---TACACCTACCTCCCGGAGACCCGCGGCCGGACGCTG 1529
Db        416 CysPheAlaSerIleTrpPheValGlnLysLysValPheGluThrArgAsnArgSerLeu 435

Qy       1530 GAGGAGATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAGCC 1589
Db        436 GluAspIle--------------GluAlaThrLeuArgAlaLysThrGlyGluAspAla 450

Qy       1590 AAGGAG 1595
Db        451 AlaGlu 452

RESULT 6
ARAE_ECOLI
ID   ARAE_ECOLI     STANDARD;      PRT;   472 AA.
AC   P09830; Q46937;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-MAR-1989 (Rel. 10, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Arabinose-proton symporter (Arabinose transporter).
GN   ARAE OR B2841 OR Z4161 OR ECS3698.
OS   Escherichia coli, and
OS   Escherichia coli O157:H7.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562, 83334;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87115869; PubMed=3543693;
RA   Maiden M.C.J., Davis E.O., Baldwin S.A., Moore D.C.M.,
```

```
RA   Henderson P.J.F.;
RT   "Mammalian and bacterial sugar transport proteins are homologous.";
RL   Nature 325:641-643(1987).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / JM2433;
RX   MEDLINE=88228015; PubMed=2836407;
RA   Maiden M.C.J., Jones-Mortimer M.C., Henderson P.J.F.;
RT   "The cloning, DNA sequence, and overexpression of the gene araE
RT   coding for arabinose-proton symport in Escherichia coli K12.";
RL   J. Biol. Chem. 263:8003-8010(1988).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX   MEDLINE=21074935; PubMed=11206551;
RA   Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA   Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA   Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA   Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA   Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA   Welch R.A., Blattner F.R.;
RT   "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL   Nature 409:529-533(2001).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / RIMD 0509952;
RX   MEDLINE=21156231; PubMed=11258796;
RA   Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA   Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA   Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA   Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT   "Complete genome sequence of enterohemorrhagic Escherichia coli
RT   O157:H7 and genomic comparison with a laboratory strain K-12.";
RL   DNA Res. 8:11-22(2001).
RN   [6]
RP   PRELIMINARY SEQUENCE OF 1-28 FROM N.A.
RX   MEDLINE=84114868; PubMed=6319708;
RA   Stoner C., Schleif R.F.;
RT   "The araE low affinity L-arabinose transport promoter. Cloning,
RT   sequence, transcription start site and DNA binding sites of
RT   regulatory proteins.";
RL   J. Mol. Biol. 171:369-381(1983).
CC   -!- FUNCTION: Uptake of arabinose across the boundary membrane with
CC       the concomitant export of a proton (symport system).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC   -!- SIMILARITY: Belongs to the sugar transporter family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; J03732; AAA23469.1; -.
DR   EMBL; X00272; CAA25075.1; -.
DR   EMBL; U29581; AAB40488.1; -.
DR   EMBL; AE000368; AAC75880.1; -.
DR   EMBL; AE005513; AAG57953.1; -.
DR   EMBL; AP002563; BAB37121.1; -.
DR   PIR; B26430; B26430.
DR   PIR; B91091; B91091.
DR   PIR; E85936; E85936.
DR   EcoGene; EG10056; araE.
DR   InterPro; IPR007114; MFS.
DR   InterPro; IPR005828; Sub_transporter.
DR   InterPro; IPR005829; Sug_transporter.
DR   InterPro; IPR003663; Sugar_transpt.
DR   Pfam; PF00083; sugar_tr; 1.
DR   PRINTS; PR00171; SUGRTRNSPORT.
DR   TIGRFAMs; TIGR00879; SP; 1.
DR   PROSITE; PS50850; MFS; 1.
DR   PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR   PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW   Transport; Sugar transport; Transmembrane; Inner membrane; Symport;
KW   Complete proteome.
FT   DOMAIN        1     29       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     30     50       1 (POTENTIAL).
FT   DOMAIN       51     63       PERIPLASMIC (POTENTIAL).
FT   TRANSMEM     64     84       2 (POTENTIAL).
FT   DOMAIN       85     91       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     92    112       3 (POTENTIAL).
FT   DOMAIN      113    114       PERIPLASMIC (POTENTIAL).
FT   TRANSMEM    115    135       4 (POTENTIAL).
FT   DOMAIN      136    154       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    155    175       5 (POTENTIAL).
FT   DOMAIN      176    178       PERIPLASMIC (POTENTIAL).
FT   TRANSMEM    179    199       6 (POTENTIAL).
FT   DOMAIN      200    257       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    258    278       7 (POTENTIAL).
FT   DOMAIN      279    297       PERIPLASMIC (POTENTIAL).
FT   TRANSMEM    298    318       8 (POTENTIAL).
FT   DOMAIN      319    325       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    326    346       9 (POTENTIAL).
FT   DOMAIN      347    361       PERIPLASMIC (POTENTIAL).
FT   TRANSMEM    362    382       10 (POTENTIAL).
FT   DOMAIN      383    404       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    405    425       11 (POTENTIAL).
FT   DOMAIN      426    427       PERIPLASMIC (POTENTIAL).
FT   TRANSMEM    428    446       12 (POTENTIAL).
FT   DOMAIN      447    472       CYTOPLASMIC (POTENTIAL).
FT   CONFLICT     26     28       SVA -> YDR (IN REF. 6).
SQ   SEQUENCE   472 AA;  51684 MW;  411990A441D44393 CRC64;

Alignment Scores:
Pred. No.:                6.77e-21        Length:          472
Score:                    579.50          Matches:         146
Percent Similarity:       51.22%          Conservative:    105
Best Local Similarity:    29.80%          Mismatches:      192
Query Match:              16.03%          Indels:          47
DB:                       1               Gaps:            11

US-10-051-902A-21 (1-2017) x ARAE_ECOLI (1-472)

Qy        102 TCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCCTTC 161
              ||||||       |||   ::::::          ::::::   |||:::       :::
Db          8 SerAlaLeuThrProArgSerLeuArgAspThrArgArgMetAsnMetPheValSerVal 27

Qy        162 GCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTGATG 221
              |||   |||:::                  :::|||   |||   ||||||||||||:::
Db         28 AlaAlaAlaVal------------AlaGlyLeuLeuPheGlyLeuAspIleGlyValIle 43

Qy        222 AGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTT 281
              :::||||||      :::|||          |||   ::::::            |||
Db         44 AlaGlyAlaLeuProPheIleThrAspHisPheValLeuThrSerArgLeuGlnGluTrp 63

Qy        282 CTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGGCGGACGTCG 341
              ::::::      :::   |||   :::   ||||||:::          |||       |||
Db         64 ValValSerSerMetMetLeuGlyAlaAlaIleGlyAlaLeuPheAsnGlyTrpLeuSer 83

Qy        342 GACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGSG 401
                 :::||||||:::|||:::::::::      |||:::::|||       |||
Db         84 PheArgLeuGlyArgLysTyrSerLeuMetAlaGlyAlaIleLeuPheValLeuGlySer 103

Qy        402 TTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGC 461
```

```
Db        104 IleGlySerAlaPheAlaThrSerValGluMetLeuIleAlaAlaArgValValLeuGly 123

Qy        462 ATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCG 521
              |||   ||||||   |||         ||||||:::|||   :::|||::::::
Db        124 IleAlaValGlyIleAlaSerTyrThrAlaProLeuTyrLeuSerGluMetAlaSerGlu 143

Qy        522 TCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTC 581
              :::   ||||||   :::   |||      ::::::   :::      ||||||:::|||
Db        144 AsnValArgGlyLysMetIleSerMetTyrGlnLeuMetValThrLeuGlyIleValLeu 163

Qy        582 GGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATG 641
                 ::::::|||:::   |||||||||          :::      ||||||   |||
Db        164 AlaPheLeuSerAspThrAlaPheSer---------TyrSerGlyAsnTrpArgAlaMet 180

Qy        642 CTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAG 701
              |||||:::   |||   |||:::|||||||||   ::::::|||:::   :::|||
Db        181 LeuGlyValLeuAlaLeuProAlaValLeuLeuIleIleLeuValValPheLeuProAsn 200

Qy        702 TCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAG 761
              |||||||||||||||      |||||||||      :::|||:::   ||||||
Db        201 SerProArgTrpLeuAlaGluLysGlyArgHisIleGluAlaGluGluValLeuArgMet 220

Qy        762 ACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGC 821
                  ||||||:::|||:::|||   |||   |||   :::|||:::   :::
Db        221 LeuArgAspThrSerGluLysAlaArgGluGluLeuAsnGluIleArgGluSerLeuLys 240

Qy        822 ATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAG 881
              :::                                  |||:::|||
Db        241 Leu-----------------------------------LysGlnGlyGly--------- 245

Qy        882 AAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTG 941
                     |||         :::                  :::||||||   :::
Db        246 ---------Trp------AlaLeuPheLysIleAsnArgAsnValArgArgAlaValPhe 260

Qy        942 TCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTAC 1001
                 |||:::   :::         |||         |||::::::   :::::::::|||
Db        261 LeuGlyMetLeuLeuGlnAlaMetGlnGlnPheThrGlyMetAsnIleIleMetTyrTyr 280

Qy       1002 AGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGCACCACT 1061
              :::|||   :::||||||      |||   |||      :::      :::      |||
Db        281 AlaProArgIlePheLysMetAlaGlyPheThrThrThrGluGlnGlnMetIleAlaThr 300

Qy       1062 TGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGC 1121
                   |||:::|||                  :::|||   |||   :::|||
Db        301 LeuValValGlyLeuThrPheMetPheAlaThrPheIleAlaValPheThrValAspLys 320

Qy       1122 GTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGC 1181
              ||||||:::|||   |||                |||      :::::::::|||
Db        321 AlaGlyArgLysProAlaLeu------LysIleGlyPheSerValMetAlaLeu------ 336

Qy       1182 CTCGGCGCCGGGCTCACCGTCGTCGGCCAG----------------------CACCCC 1217
                    |||      |||:::|||
Db        337 ---------GlyThrLeuValLeuGlyTyrCysLeuMetGlnPheAspAsnGlyThrAla 353

Qy       1218 GACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCC 1277
              :::   :::   |||         ||||||:::      |||:::      :::|||
Db        354 SerSerGlyLeuSerTrp---------LeuSerValGlyMetThrMetMetCysIleAla 370

Qy       1278 TTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTC 1337
                 :::::::::         |||:::   |||:::      |||||||||   ||||||
Db        371 GlyTyrAlaMetSerAlaAlaProValValTrpIleLeuCysSerGluIleGlnProLeu 390

Qy       1338 CAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTC 1397
              :::   |||      |||   :::                  |||   |||::::::   :::
Db        391 LysCysArgAspPheGlyIleThrCysSerThrThrThrAsnTrpValSerAsnMetIle 410

Qy       1398 ATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTC 1457
              |||      |||||||||:::|||      :::|||         |||:::|||:::|||
Db        411 IleGlyAlaThrPheLeuThrLeuLeuAspSerIleGlyAlaAlaGlyThrPheTrpLeu 430

Qy       1458 TACTCCGGCATC---GCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACC 1514
              |||:::   :::      |||         :::   |||:::      :::|||||||||
```

```
Db        431 TyrThrAlaLeuAsnIleAlaPheValGlyIleThrPheTrp---LeuIleProGluThr 449

Qy       1515 CGCGGCCGGACGCTGGAGGAGATGAGCAAG 1544
                   :::      |||||||||   :::   :::
Db        450 LysAsnValThrLeuGluHisIleGluArg 459
```

```
RESULT 7
ITR1_YEAST
ID   ITR1_YEAST     STANDARD;      PRT;   584 AA.
AC   P30605;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Myo-inositol transporter 1.
GN   ITR1 OR YDR497C OR D9719.3.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91250431; PubMed=2040626;
RA   Nikawa J.-I., Tsukagoshi Y., Yamashita S.;
RT   "Isolation and characterization of two distinct myo-inositol
RT   transporter genes of Saccharomyces cerevisiae.";
RL   J. Biol. Chem. 266:11184-11191(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Dietrich F.S., Mulligan J., Allen E., Araujo R., Aviles E.,
RA   Berno A., Carpenter J., Chen E., Cherry J.M., Chung E., Duncan M.,
RA   Hunicke-Smith S., Hyman R., Komp C., Lashkari D., Lew H., Lin D.,
RA   Mosedale D., Nakahara K., Namath A., Oefner P., Oh C., Petel F.X.,
RA   Roberts D., Schramm S., Schroeder M., Shogren T., Shroff N.,
RA   Winant A., Yelton M., Botstein D., Davis R.W.;
RL   Submitted (AUG-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: MAJOR TRANSPORTER FOR MYO-INOSITOL.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to the sugar transporter family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; D90352; BAA14366.1; -.
DR   EMBL; U33057; AAB64939.1; -.
DR   PIR; S69555; S69555.
DR   GermOnline; 140989; -.
DR   SGD; S0002905; ITR1.
DR   GO; GO:0005365; F:myo-inositol transporter activity; IMP.
DR   GO; GO:0015798; P:myo-inositol transport; IMP.
DR   InterPro; IPR007114; MFS.
DR   InterPro; IPR005828; Sub_transporter.
DR   InterPro; IPR005829; Sug_transporter.
DR   InterPro; IPR003663; Sugar_transpt.
DR   Pfam; PF00083; sugar_tr; 1.
DR   PRINTS; PR00171; SUGRTRNSPORT.
DR   TIGRFAMs; TIGR00879; SP; 1.
DR   PROSITE; PS50850; MFS; 1.
DR   PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.
DR   PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW   Transmembrane; Sugar transport; Glycoprotein.
FT   DOMAIN        1     81       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     82    102       1 (POTENTIAL).
FT   DOMAIN      103    129       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    130    150       2 (POTENTIAL).
FT   DOMAIN      151    163       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    164    184       3 (POTENTIAL).
FT   DOMAIN      185    186       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    187    207       4 (POTENTIAL).
```

```
FT   DOMAIN       208    215       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     216    236       5 (POTENTIAL).
FT   DOMAIN       237    246       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     247    267       6 (POTENTIAL).
FT   DOMAIN       268    349       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     350    370       7 (POTENTIAL).
FT   DOMAIN       371    376       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     377    397       8 (POTENTIAL).
FT   DOMAIN       398    400       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     401    421       9 (POTENTIAL).
FT   DOMAIN       422    441       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     442    462       10 (POTENTIAL).
FT   DOMAIN       463    486       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     487    507       11 (POTENTIAL).
FT   DOMAIN       508    510       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     511    531       12 (POTENTIAL).
FT   DOMAIN       532    584       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     371    371       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT      43     44       TL -> HI (IN REF. 1).
SQ   SEQUENCE   584 AA;  63569 MW;  42543E30A102DC65 CRC64;

Alignment Scores:
Pred. No.:                1.01e-20      Length:          584
Score:                    576.00        Matches:         148
Percent Similarity:       50.40%        Conservative:    106
Best Local Similarity:    29.37%        Mismatches:      206
Query Match:              15.94%        Indels:          44
DB:                       1             Gaps:            11

US-10-051-902A-21 (1-2017) x ITR1_YEAST (1-584)

Qy        153 TTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATC 212
Db         85 PheIleIleThrLeuThrPheValAlaSerIleSerGlyPheMetPheGlyTyrAspThr 104

Qy        213 GGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAAC---------ATCAGT 263
Db        105 GlyTyrIleSerSerAlaLeuIleSerIleGlyThrAspLeuAspHisLysValLeuThr 124

Qy        264 GACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTC 323
Db        125 TyrGlyGluLysGluIleValThrAlaAlaThrSerLeuGlyAlaLeuIleThrSerIle 144

Qy        324 GCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTC 383
Db        145 PheAlaGlyThrAlaAlaAspIlePheGlyArgLysArgCysLeuMetGlySerAsnLeu 164

Qy        384 ATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTC 443
Db        165 MetPheValIleGlyAlaIleLeuGlnValSerAlaHisThrPheTrpGlnMetAlaVal 184

Qy        444 GGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACC 503
Db        185 GlyArgLeuIleMetGlyPheGlyValGlyIleGlySerLeuIleAlaProLeuPheIle 204

Qy        504 GCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATC 563
Db        205 SerGluIleAlaProLysMetIleArgGlyArgLeuThrValIleAsnSerLeuTrpLeu 224

Qy        564 AACTTCGGCATCCTGCTCGGGTACGTCTCG------------AACTATGCTTTCTCCCGC 611
Db        225 ThrGlyGlyGlnLeuValAlaTyrGlyCysGlyAlaGlyLeuAsnTyr------------ 240

Qy        612 TTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTG 671
Db        241 ------ValAsnAsnGlyTrpArgIleLeuValGlyLeuSerLeuIleProThrAlaVal 258

Qy        672 CTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGC 731
Db        259 GlnPheThrCysLeuCysPheLeuProAspThrProArgTyrTyrValMetLysGlyAsp 278

Qy        732 CTCGCGGACGCCAAGGTGGTGCTGGAGAAGACC---TCCGACACGGCGGAGGAGGCCGCG 788
```

```
Db        279 LeuAlaArgAlaThrGluValLeuLysArgSerTyrThrAspThrSerGluGluIleIle 298

Qy        789 GAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTG 848
Db        299 GluArgLysValGlu----------------------------------------GluLeu 305

Qy        849 GTGACCGTCCCCAAGAGA------GGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTC 902
Db        306 ValThrLeuAsnGlnSerIleProGlyLysAsnValProGluLysValTrpAsnThrIle 325

Qy        903 ---ATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCAC 959
Db        326 LysGluLeuHisThrValProSerAsnLeuArgAlaLeuIleIleGlyCysGlyLeuGln 345

Qy        960 TTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAG 1019
Db        346 AlaIleGlnGlnPheThrGlyTrpAsnSerLeuMetTyrPheSerGlyThrIlePheGlu 365

Qy       1020 AGCCCCGGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACC 1079
Db        366 ThrValGlyPheLysAsnSerSer---------AlaValSerIleIleValSerGlyThr 382

Qy       1080 AAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTG 1139
Db        383 AsnPheIlePheThrLeuValAlaPhePheSerIleAspLysIleGlyArgArgThrIle 402

Qy       1140 TTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACC 1199
Db        403 LeuLeuIleGlyLeuProGlyMetThrMetAlaLeuValValCysSerIleAlaPheHis 422

Qy       1200 GTCGTCGGCCAGCACCCCGACGCCAAGATA--------------------------CCT 1232
Db        423 PheLeuGlyIleLysPheAspGlyAlaValAlaValValValSerSerGlyPheSerSer 442

Qy       1233 TGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGC 1292
Db        443 TrpGlyIle---ValIleIleValPheIleIleValPheAlaAlaPheTyrAlaLeuGly 461

Qy       1293 CTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTG 1352
Db        462 IleGlyThrValProTrp---GlnGlnSerGluLeuPheProGlnAsnValArgGlyIle 480

Qy       1353 GGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTC 1412
Db        481 GlyThrSerTyrAlaThrAlaThrAsnTrpAlaGlySerLeuValIleAlaSerThrPhe 500

Qy       1413 CTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCC 1472
Db        501 LeuThrMetLeuGlnAsnIleThrProAlaGlyThrPheAlaPhePheAlaGlyLeuSer 520

Qy       1473 GCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAG 1532
Db        521 CysLeuSerThrIlePheCysTyrPheCysTyrProGluLeuSerGlyLeuGluLeuGlu 540

Qy       1533 GAGATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAGCCAAG 1592
Db        541 GluValGlnThrIleLeuLysAspGlyPheAsnIleLysAlaSerLysAlaLeuAlaLys 560

Qy       1593 GAGAAGAAGAAG 1604
Db        561 LysArgLysGln 564

RESULT 8
XYLE_ECOLI
ID   XYLE_ECOLI     STANDARD;      PRT;   491 AA.
AC   P09098;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-MAR-1989 (Rel. 10, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   D-xylose-proton symporter (D-xylose transporter).
GN   XYLE OR B4031 OR Z5629 OR ECS5014.
OS   Escherichia coli, and
OS   Escherichia coli O157:H7.
```

```
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562, 83334;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=88007632; PubMed=2820984;
RA   Davis E.O., Henderson P.J.F.;
RT   "The cloning and DNA sequence of the gene xylE for xylose-proton
RT   symport in Escherichia coli K12.";
RL   J. Biol. Chem. 262:13928-13932(1987).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87115869; PubMed=3543693;
RA   Maiden M.C.J., Davis E.O., Baldwin S.A., Moore D.C.M.,
RA   Henderson P.J.F.;
RT   "Mammalian and bacterial sugar transport proteins are homologous.";
RL   Nature 325:641-643(1987).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=94089392; PubMed=8265357;
RA   Blattner F.R., Burland V.D., Plunkett G. III, Sofia H.J.,
RA   Daniels D.L.;
RT   "Analysis of the Escherichia coli genome. IV. DNA sequence of the
RT   region from 89.2 to 92.8 minutes.";
RL   Nucleic Acids Res. 21:5408-5417(1993).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX   MEDLINE=21074935; PubMed=11206551;
RA   Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA   Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA   Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA   Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA   Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA   Welch R.A., Blattner F.R.;
RT   "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL   Nature 409:529-533(2001).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / RIMD 0509952;
RX   MEDLINE=21156231; PubMed=11258796;
RA   Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA   Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA   Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA   Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT   "Complete genome sequence of enterohemorrhagic Escherichia coli
RT   O157:H7 and genomic comparison with a laboratory strain K-12.";
RL   DNA Res. 8:11-22(2001).
RN   [6]
RP   SEQUENCE OF 1-192 FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=88234001; PubMed=2836810;
RA   Francoz E., Dassa E.;
RT   "3' end of the malEFG operon in E.coli: localization of the
RT   transcription termination site.";
RL   Nucleic Acids Res. 16:4097-4109(1988).
CC   -!- FUNCTION: Uptake of D-xylose across the boundary membrane with the
CC       concomitant export of a proton (symport system).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC   -!- INDUCTION: By xylose.
CC   -!- MISCELLANEOUS: E.COLI HAS TWO D-XYLOSE TRANSPORT SYSTEMS THAT
CC       ACCUMULATE SUGAR AGAINST A CONCENTRATION GRADIENT: THE XYLE SYSTEM
CC       WHICH UTILIZES THE ELECTROCHEMICAL GRADIENT OF PROTONS AND THAT IS
CC       INSENSITIVE TO COLD OSMOTIC SHOCK AND THE XYLF SYSTEM THAT USES A
CC       HIGH-ENERGY PHOSPHATE COMPOUND AND IS SENSITIVE TO COLD OSMOTIC
CC       SHOCK.
CC   -!- SIMILARITY: Belongs to the sugar transporter family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; J02812; AAA79016.1; -.
DR   EMBL; U00006; AAC43125.1; -.
DR   EMBL; AE000476; AAC77001.1; -.
DR   EMBL; AE005636; AAG59230.1; -.
DR   EMBL; AP002568; BAB38437.1; -.
DR   EMBL; X06663; CAA29863.1; -.
DR   PIR; A26430; A26430.
DR   PIR; B86096; B86096.
DR   PIR; F91255; F91255.
DR   EcoGene; EG11076; xylE.
DR   InterPro; IPR007114; MFS.
DR   InterPro; IPR005828; Sub_transporter.
DR   InterPro; IPR005829; Sug_transporter.
DR   InterPro; IPR003663; Sugar_transpt.
DR   Pfam; PF00083; sugar_tr; 1.
DR   PRINTS; PR00171; SUGRTRNSPORT.
DR   TIGRFAMs; TIGR00879; SP; 1.
DR   PROSITE; PS50850; MFS; 1.
DR   PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR   PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW   Transport; Sugar transport; Transmembrane; Inner membrane; Symport;
KW   Complete proteome.
FT   DOMAIN        1      9       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     10     30       1 (POTENTIAL).
FT   DOMAIN       31     55       PERIPLASMIC (POTENTIAL).
FT   TRANSMEM     56     76       2 (POTENTIAL).
FT   DOMAIN       77     89       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     90    110       3 (POTENTIAL).
FT   DOMAIN      111    133       PERIPLASMIC (POTENTIAL).
FT   TRANSMEM    134    154       4 (POTENTIAL).
FT   DOMAIN      155    165       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    166    186       5 (POTENTIAL).
FT   DOMAIN      187    200       PERIPLASMIC (POTENTIAL).
FT   TRANSMEM    201    221       6 (POTENTIAL).
FT   DOMAIN      222    272       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    273    293       7 (POTENTIAL).
FT   DOMAIN      294    312       PERIPLASMIC (POTENTIAL).
FT   TRANSMEM    313    333       8 (POTENTIAL).
FT   DOMAIN      334    343       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    344    364       9 (POTENTIAL).
FT   DOMAIN      365    369       PERIPLASMIC (POTENTIAL).
FT   TRANSMEM    370    390       10 (POTENTIAL).
FT   DOMAIN      391    407       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    408    428       11 (POTENTIAL).
FT   DOMAIN      429    442       PERIPLASMIC (POTENTIAL).
FT   TRANSMEM    443    463       12 (POTENTIAL).
FT   DOMAIN      464    491       CYTOPLASMIC (POTENTIAL).
FT   CONFLICT     64     64       A -> V (IN REF. 6).
SQ   SEQUENCE   491 AA;  53608 MW;  2AF1AF9756C0B722 CRC64;

Alignment Scores:
Pred. No.:                1.24e-20        Length:          491
Score:                    574.00          Matches:         163
Percent Similarity:       46.49%          Conservative:    89
Best Local Similarity:    30.07%          Mismatches:      184
Query Match:              15.88%          Indels:          106
DB:                       1               Gaps:            17

US-10-051-902A-21 (1-2017) x XYLE_ECOLI (1-491)

Qy        144 AACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGC 203
              |||      :::   |||:::      ::::::|||::::::      :::|||   |||
Db          6 AsnSerSerTyrIlePheSerIleThrLeuValAlaThrLeuGlyGlyLeuLeuPheGly 25

Qy        204 TACGATATCGGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTC--------- 254
              ||||||      |||:::||||||           :::        |||
Db         26 TyrAspThrAlaValIleSerGlyThrValGluSerLeuAsnThrValPheValAlaPro 45
```

```
Qy      255 ---AACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTC 311
Db       46 GlnAsnLeuSerGluSerAlaAlaAsnSerLeuLeuGlyPheCysValAlaSerAlaLeu 65

Qy      312 ATCGGCTCCTTCGCGGCGGGGCGG-----------ACGTCGGACTGGATCGGCCGGCGG 359
Db       66 IleGlyCysIleIleGlyGlyAlaLeuGlyGlyTyrCysSerAsnArgPheGlyArgArg 85

Qy      360 TACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGG---------------------- 398
Db       86 AspSerLeuLysIleAlaAlaValLeuPhePheIleSerGlyValGlySerAlaTrpPro 105

Qy      399 ----------------------------------------GSGTTCCTCATGGGGTTC 416
Db      106 GluLeuGlyPheThrSerIleAsnProAspAsnThrValProValTyrLeuAlaGlyTyr 125

Qy      417 GCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTAC 476
Db      126 ValProGluPheValIleTyr---------ArgIleIleGlyGlyIleGlyValGlyLeu 142

Qy      477 GCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTC 536
Db      143 AlaSerMetLeuSerProMetTyrIleAlaGluLeuAlaProAlaHisIleArgGlyLys 162

Qy      537 CTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGGTACGTCTCGAAC 596
Db      163 LeuValSerPheAsnGlnPheAlaIleIlePheGlyGlnLeuLeuValTyrCysValAsn 182

Qy      597 TATGCTTTCTCCCGCTTGCCG-----------CTGAACCTC---GGGTGGCGCATCATG 641
Db      183 TyrPheIleAlaArgSerGlyAspAlaSerTrpLeuAsnThrAspGlyTrpArgTyrMet 202

Qy      642 CTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAG 701
Db      203 PheAlaSerGluCysIleProAlaLeuLeuPheLeuMetLeuLeuTyrThrValProGlu 222

Qy      702 TCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAG 761
Db      223 SerProArgTrpLeuMetSerArgGlyLysGlnGluGlnAlaGluGlyIleLeuArgLys 242

Qy      762 ACC------TCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCC 815
Db      243 IleMetGlyAsnThrLeuAlaThrGlnAlaValGlnGlu---------------------- 255

Qy      816 GCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGA 875
Db      256 ------IleLysHisSerLeuAspHisGly-----------------ArgLysThrGly 267

Qy      876 AACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATC 935
Db      268 Gly-------------------------------------------------------ArgLeu 270

Qy      936 CTGCTGTCCGGGATCGGCATC--------------------CACTTCTTCCAGCATGCG 974
Db      271 LeuMetPheGlyValGlyValIleValIleGlyValMetLeuSerIlePheGlnGlnPhe 290

Qy      975 TTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACG 1034
Db      291 ValGlyIleAsnValValLeuTyrTyrAlaProGluValPheLysThrLeuGlyAlaSer 310

Qy     1035 AACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATC 1094
Db      311 ThrAspIleAlaLeuLeuGlnThrIle---IleValGlyValIleAsnLeuThrPheThr 329

Qy     1095 TTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACG 1154
Db      330 ValLeuAlaIleMetThrValAspLysPheGlyArgLysProLeuGln------------ 345

Qy     1155 GGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCAC 1214
Db      346 ---------IleIleGlyAlaLeuGlyMetAlaIleGlyMetPheSerLeuGlyThrAla 362

Qy     1215 CCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTC 1274
Db      363 PheTyrThrGlnAlaProGlyIleValAlaLeu------LeuSerMetLeuPheTyrVal 380

Qy     1275 GCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCG 1334
Db      381 AlaAlaPheAlaMetSerTrpGlyProValCysTrpValLeuLeuSerGluIlePhePro 400

Qy     1335 CTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGC 1394
Db      401 AsnAlaIleArgGlyLysAlaLeuAlaIleAlaValAlaAlaGlnTrpLeuAlaAsnTyr 420

Qy     1395 GTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATC-------------- 1439
Db      421 PheValSerTrpThrPheProMetMetAspLysAsnSerTrpLeuValAlaHisPheHis 440

Qy     1440 ---GGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTAC 1496
Db      441 AsnGlyPheSerTyrTrpIleTyrGlyCysMetGlyValLeuAlaAlaLeuPheMetTrp 460

Qy     1497 ACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGAC 1556
Db      461 LysPheValProGluThrLysGlyLysThrLeuGluGluLeuGluAlaLeuTrp------ 478

Qy     1557 ACGGCCGCCGCCTCGGAATCAGACGAGCCAGCCAAGGAGAAGAAGAAGGTGGAAATGGCC 1616
Db      479 ------------------------GluPro------GluThrLysLysThrGlnGlnThr 488

Qy     1617 GCCACT 1622
Db      489 AlaThr 490
```

RESULT 9
GALP_ECOLI
ID   GALP_ECOLI   STANDARD;   PRT;   464 AA.
AC   P37021;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Galactose-proton symporter (Galactose transporter).
GN   GALP OR B2943 OR C3529.
OS   Escherichia coli, and
OS   Escherichia coli O6.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562, 217992;
RN   [1]
RP   SEQUENCE FROM N.A., AND CHARACTERIZATION.
RA   Roberts P.E.;
RL   Thesis (1992), University of Cambridge, U.K.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX   MEDLINE=22388234; PubMed=12471157;
RA   Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA   Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA   Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA   Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT   "Extensive mosaic structure revealed by the complete genome sequence
RT   of uropathogenic Escherichia coli.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
CC   -!- FUNCTION: UPTAKE OF GALACTOSE ACROSS THE BOUNDARY MEMBRANE WITH
CC   THE CONCOMITANT EXPORT OF A PROTON (SYMPORT SYSTEM).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.

```
CC   -!- SIMILARITY: Belongs to the sugar transporter family.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; U28377; AAA69110.1; -.
DR   EMBL; AE000377; AAC75980.1; -.
DR   EMBL; AE016766; AAN81977.1; ALT_INIT.
DR   PIR; F65079; F65079.
DR   EcoGene; EG12148; galP.
DR   InterPro; IPR007114; MFS.
DR   InterPro; IPR005828; Sub_transporter.
DR   InterPro; IPR005829; Sug_transporter.
DR   InterPro; IPR003663; Sugar_transpt.
DR   Pfam; PF00083; sugar_tr; 1.
DR   PRINTS; PR00171; SUGRTRNSPORT.
DR   TIGRFAMs; TIGR00879; SP; 1.
DR   PROSITE; PS50850; MFS; 1.
DR   PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR   PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW   Transport; Sugar transport; Transmembrane; Inner membrane; Symport;
KW   Complete proteome.
FT   DOMAIN        1     15       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     16     36       1 (POTENTIAL).
FT   DOMAIN       37     56       PERIPLASMIC (POTENTIAL).
FT   TRANSMEM     57     77       2 (POTENTIAL).
FT   DOMAIN       78     84       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     85    105       3 (POTENTIAL).
FT   DOMAIN      106    112       PERIPLASMIC (POTENTIAL).
FT   TRANSMEM    113    133       4 (POTENTIAL).
FT   DOMAIN      134    139       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    140    160       5 (POTENTIAL).
FT   DOMAIN      161    171       PERIPLASMIC (POTENTIAL).
FT   TRANSMEM    172    192       6 (POTENTIAL).
FT   DOMAIN      193    250       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    251    271       7 (POTENTIAL).
FT   DOMAIN      272    290       PERIPLASMIC (POTENTIAL).
FT   TRANSMEM    291    311       8 (POTENTIAL).
FT   DOMAIN      312    321       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    322    342       9 (POTENTIAL).
FT   DOMAIN      343    351       PERIPLASMIC (POTENTIAL).
FT   TRANSMEM    352    372       10 (POTENTIAL).
FT   DOMAIN      373    394       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    395    415       11 (POTENTIAL).
FT   DOMAIN      416    416       PERIPLASMIC (POTENTIAL).
FT   TRANSMEM    417    437       12 (POTENTIAL).
FT   DOMAIN      438    464       CYTOPLASMIC (POTENTIAL).
SQ   SEQUENCE   464 AA;  50982 MW;  07E08935BD8E3F8E CRC64;

Alignment Scores:
Pred. No.:              1.46e-20      Length:        464
Score:                  572.50        Matches:       141
Percent Similarity:     49.79%        Conservative:  99
Best Local Similarity:  29.25%        Mismatches:    205
Query Match:            15.84%        Indels:        37
DB:                     1             Gaps:          8

US-10-051-902A-21 (1-2017) x GALP_ECOLI (1-464)

Qy        111 CTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGC------AACGTCCGGTTCGCCTTCGCC 164
                :::|||:::|||          ||||||:::|||       |||               |||
Db          1 MetProAspAla---------LysLysGlnGlyArgSerAsnLysAlaMetThrPhePhe 17

Qy        165 TGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTGATGAGC 224
                          ||||||::::::      :::|||   |||   ||||||||||||::::::
Db         18 ValCysPheLeuAlaAlaLeuAlaGlyLeuLeuPheGlyLeuAspIleGlyValIleAla 37

Qy        225 GGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTC 284
              ||||||        :::|||       :::|||   |||:::           |||    :::
Db         38 GlyAlaLeuProPheIleAlaAspGluPheGlnIleThrSerHisThrGlnGluTrpVal 57

Qy        285 ATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGAC 344
              :::       :::          :::   :::|||:::       :::|||       |||
Db         58 ValSerSerMetMetPheGlyAlaAlaValGlyAlaValGlySerGlyTrpLeuSerPhe 77

Qy        345 TGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGSGTTC 404
                 :::||||||:::    :::::::::          |||::::::|||   ||||||
Db         78 LysLeuGlyArgLysLysSerLeuMetIleGlyAlaIleLeuPheValAlaGlySerLeu 97

Qy        405 CTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATC 464
                       |||   |||      :::|||:::       |||    :::    |||:::
Db         98 PheSerAlaAlaAlaProAsnValGluValLeuIleLeuSerArgValLeuLeuGlyLeu 117

Qy        465 GGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCG 524
                 ||||||   |||         ||||||:::|||    :::|||::::::|||
Db        118 AlaValGlyValAlaSerTyrThrAlaProLeuTyrLeuSerGluIleAlaProGluLys 137

Qy        525 GCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGG 584
                 ||||||   :::      |||       ::::::   |||   |||||||||
Db        138 IleArgGlySerMetIleSerMetTyrGlnLeuMetIleThrIleGlyIleLeuGlyAla 157

Qy        585 TACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTC 644
              |||:::|||:::   |||||||||                      ||||||   ||||||
Db        158 TyrLeuSerAspThrAlaPheSer---------TyrThrGlyAlaTrpArgTrpMetLeu 174

Qy        645 GGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCG 704
              |||:::       |||::::::||||||   :::    |||       :::|||:::|||
Db        175 GlyValIleIleIleProAlaIleLeuLeuLeuIleGlyValPhePheLeuProAspSer 194

Qy        705 CCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACC 764
              |||||||||          |||   |||       ||||||:::       ||||||  :::
Db        195 ProArgTrpPheAlaAlaLysArgArgPheValAspAlaGluArgValLeuLeuArgLeu 214

Qy        765 TCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATC 824
                 ||||||:::   ||||||       |||   :::|||:::   :::          :::
Db        215 ArgAspThrSerAlaGluAlaLysArgGluLeuAspGluIleArgGluSerLeuGlnVal 234

Qy        825 CCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAG 884
                 :::                                         ||||||
Db        235 LysGln---------------------------------------SerGly--------- 238

Qy        885 CGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCC 944
                    |||      :::                          ||||||    :::
Db        239 ------Trp------AlaLeuPheLysGluAsnSerAsnPheArgArgAlaValPheLeu 254

Qy        945 GGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGC 1004
              |||:::   :::          |||       |||::::::   :::::::::|||:::
Db        255 GlyValLeuLeuGlnValMetGlnGlnPheThrGlyMetAsnValIleMetTyrTyrAla 274

Qy       1005 CCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGCACCACTTGG 1064
              |||   :::|||:::       |||   ||||||             :::      |||
Db        275 ProLysIlePheGluLeuAlaGlyTyrThrAsnThrThrGluGlnMetTrpGlyThrVal 294

Qy       1065 CCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTC 1124
                    |||:::|||       |||          :::|||          :::|||
Db        295 IleValGlyLeuThrAsnValLeuAlaThrPheIleAlaIleGlyLeuValAspArgTrp 314

Qy       1125 GGCGCGGCGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTC 1184
              ||||||:::|||   |||                       |||       :::    :::
Db        315 GlyArgLysProThrLeu---------------------ThrLeuGlyPheLeuValMet 327

Qy       1185 GGCGCCGGGCTCACCGTCGTCGGC------CAGCACCCCGACGCCAAGATACCTTGGGCC 1238
                 ||||||:::   |||:::|||             |||          |||   |||
Db        328 AlaAlaGlyMetGlyValLeuGlyThrMetMetHisIleGlyIleHisSerProSerAla 347

Qy       1239 ATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGC 1298
                 :::||||||      |||   ::::::       |||::::::        |||
Db        348 GlnTyrPheAlaIleAlaMetLeuLeuMetPheIleValGlyPheAlaMetSerAlaGly 367

Qy       1299 CCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGC 1358
              |||:::   ||||||       |||||||||   ||||||:::    |||       |||
```

```
Db        368 ProLeuIleTrpValLeuCysSerGluIleGlnProLeuLysGlyArgAspPheGlyIle 387

Qy       1359 TCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCG 1418
              :::         |||   |||   :::   :::   ::::::      |||||||||:::
Db        388 ThrCysSerThrAlaThrAsnTrpIleAlaAsnMetIleValGlyAlaThrPheLeuThr 407

Qy       1419 CTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTC 1478
              :::         :::          :::|||::::::|||:::   :::     |||
Db        408 MetLeuAsnThrLeuGlyAsnAlaAsnThrPheTrpValTyrAlaAlaLeuAsnValLeu 427

Qy       1479 GCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATG 1538
                 :::               :::|||||||||:::      :::||||||   :::
Db        428 PheIleLeuLeuThrLeuTrpLeuValProGluThrLysHisValSerLeuGluHisIle 447

Qy       1539 AGCAAG 1544
                 :::
Db        448 GluArg 449
```

```
RESULT 10
ITR2_YEAST
ID   ITR2_YEAST     STANDARD;      PRT;   612 AA.
AC   P30606;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Myo-inositol transporter 2.
GN   ITR2 OR YOL103W OR HRB612.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91250431; PubMed=2040626;
RA   Nikawa J.-I., Tsukagoshi Y., Yamashita S.;
RT   "Isolation and characterization of two distinct myo-inositol
RT   transporter genes of Saccharomyces cerevisiae.";
RL   J. Biol. Chem. 266:11184-11191(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96076631; PubMed=7502582;
RA   Vandenbol M., Durand P., Portetelle D., Hilger F.;
RT   "Sequence analysis of a 44 kb DNA fragment of yeast chromosome XV
RT   including the Tyl-H3 retrotransposon, the suf1(+) frameshift
RT   suppressor gene for tRNA-Gly, the yeast transfer RNA-Thr-1a and a
RT   delta element.";
RL   Yeast 11:1069-1075(1995).
CC   -!- FUNCTION: MINOR TRANSPORTER FOR MYO-INOSITOL.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to the sugar transporter family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; D90353; BAA14367.1; -.
DR   EMBL; Z48149; CAA88159.1; -.
DR   EMBL; Z74845; CAA99119.1; -.
DR   PIR; B40538; B40538.
DR   GermOnline; 143525; -.
DR   SGD; S0005463; ITR2.
DR   GO; GO:0005365; F:myo-inositol transporter activity; IMP.
DR   GO; GO:0015798; P:myo-inositol transport; IMP.
DR   InterPro; IPR007114; MFS.
DR   InterPro; IPR005828; Sub_transporter.
DR   InterPro; IPR005829; Sug_transporter.
DR   InterPro; IPR003663; Sugar_transpt.
DR   Pfam; PF00083; sugar_tr; 1.
DR   PRINTS; PR00171; SUGRTRNSPORT.
DR   TIGRFAMs; TIGR00879; SP; 1.
DR   PROSITE; PS50850; MFS; 1.
DR   PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.
DR   PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW   Transmembrane; Sugar transport; Glycoprotein.
FT   DOMAIN        1    107       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    108    128       1 (POTENTIAL).
FT   DOMAIN      129    155       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    156    176       2 (POTENTIAL).
FT   DOMAIN      177    182       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    183    203       3 (POTENTIAL).
FT   DOMAIN      204    212       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    213    233       4 (POTENTIAL).
FT   DOMAIN      234    241       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    242    262       5 (POTENTIAL).
FT   DOMAIN      263    272       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    273    293       6 (POTENTIAL).
FT   DOMAIN      294    362       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    363    383       7 (POTENTIAL).
FT   DOMAIN      384    402       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    403    423       8 (POTENTIAL).
FT   DOMAIN      424    426       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    427    447       9 (POTENTIAL).
FT   DOMAIN      448    467       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    468    488       10 (POTENTIAL).
FT   DOMAIN      489    512       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    513    533       11 (POTENTIAL).
FT   DOMAIN      534    536       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    537    557       12 (POTENTIAL).
FT   DOMAIN      558    612       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD    397    397       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   612 AA;  67041 MW;  48E733102BAD7298 CRC64;
```

```
Alignment Scores:
Pred. No.:              1.48e-20        Length:         612
Score:                  572.50          Matches:        154
Percent Similarity:     50.00%          Conservative:   103
Best Local Similarity:  29.96%          Mismatches:     210
Query Match:            15.84%          Indels:         47
DB:                     1               Gaps:           14

US-10-051-902A-21 (1-2017) x ITR2_YEAST (1-612)

Qy        117 GAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGCGCCATCCTC 176
              ::::::::::|||                  |||           :::
Db        106 GlnSerIleSerPro--------------------PheIleIleThrLeuThrPheVal 118

Qy        177 GCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTG 236
              ||||||::::::      :::  |||||||||  |||  :::|||  |||   :::
Db        119 AlaSerIleSerGlyPheMetPheGlyTyrAspThrGlyTyrIleSerSerAlaLeuIle 138

Qy        237 TACATCAAGAAGGACTTCAAC---------ATCAGTGACGGGAAGGTGGAGGTTCTCATG 287
                 |||   :::|||   :::         ::::::   |||:::   |||::::::
Db        139 SerIleAsnArgAspLeuAspAsnLysValLeuThrTyrGlyGluLysGluLeuIleThr 158

Qy        288 GGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGG 347
                    :::|||   :::||||||   |||      ||||||      :::|||
Db        159 AlaAlaThrSerLeuGlyAlaLeuIleThrSerValGlyAlaGlyThrAlaAlaAspVal 178

Qy        348 ATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGSGTTCCTC 407
                 |||||||||      ::::::|||:::   ::::::|||      |||      |||
Db        179 PheGlyArgArgProCysLeuMetPheSerAsnLeuMetPheLeuIleGlyAlaIleLeu 198

Qy        408 ATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGC 467
                 |||      :::      :::      ||||||   :::   |||   |||
Db        199 GlnIleThrAlaHisLysPheTrpGlnMetAlaAlaGlyArgLeuIleMetGlyPheGly 218

Qy        468 GTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCG 527
              ||||||         :::|||:::|||::::::   :::|||::::::|||
Db        219 ValGlyIleGlySerLeuIleSerProLeuPheIleSerGluIleAlaProLysMetIle 238

Qy        528 CGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTAC 587
              ||||||   ||||||            :::::::::      |||   |||:::   |||
```

```
Db        239 ArgGlyArgLeuThrValIleAsnSerLeuTrpLeuThrGlyGlyGlnLeuIleAlaTyr 258

Qy        588 GTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGC 647
                      :::         :::      ||||||||||||:::::|||
Db        259 GlyCysGlyAlaGlyLeuAsnHis------ValLysAsnGlyTrpArgIleLeuValGly 276

Qy        648 ATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCG 707
              :::        |||:::||||||                 :::|||::::::|||
Db        277 LeuSerLeuIleProThrValLeuGlnPheSerPhePheCysPheLeuProAspThrPro 296

Qy        708 CGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCC 767
              |||:::   ||||||||||||   |||       ||||||:::|||||:::::::::
Db        297 ArgTyrTyrValMetLysGlyAspLeuLysArgAlaLysMetValLeuLysArgSerTyr 316

Qy        768 GACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCT 827
                    |||         :::   :::    |||   |||
Db        317 ValAsnThrGlu---------AspGluIleIleAspGlnLysVal-------------- 328

Qy        828 GAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGG 887
              |||||||||                 ::::::|||      |||                 :::
Db        329 GluGluLeuSerSerLeuAsnGlnSerIlePro------GlyLysAsnProIleThrLys 346

Qy        888 GTGTGG---------AAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTG 938
                |||         ||||||     |||         |||:::      |||   |||
Db        347 PheTrpAsnMetValLysGlu------LeuHisThrValProSerAsnPheArgAlaLeu 364

Qy        939 CTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTC 998
              :::   |||   |||:::          |||         |||   :::|||:::::::::
Db        365 IleIleGlyCysGlyLeuGlnAlaIleGlnGlnPheThrGlyTrpAsnSerLeuMetTyr 384

Qy        999 TACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGCACC 1058
              :::|||      :::|||::::::   |||       |||
Db        385 PheSerGlyThrIlePheGluThrValGlyPheLysAsnSerSer---------AlaVal 401

Qy       1059 ACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGAC 1118
              :::                |||      :::|||   |||:::|||   |||   ||||||
Db        402 SerIleIleValSerGlyThrAsnPheValPheThrLeuIleAlaPhePheCysIleAsp 421

Qy       1119 GGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATC 1178
                 :::|||||||||   :::||||||                |||:::   ::::::|||:::
Db        422 LysIleGlyArgArgTyrIleLeuLeuIleGlyLeuProGlyMetThrValAlaLeuVal 441

Qy       1179 --------------------------GGCCTC---------GGCGCCGGGCTCACCGTC 1202
                                        |||:::         ||||||      |||
Db        442 IleCysAlaIleAlaPheHisPheLeuGlyIleLysPheAsnGlyAlaAspAlaValVal 461

Qy       1203 GTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACC 1262
                        |||             |||   |||   :::   |||
Db        462 AlaSer---------AspGlyPheSerSerTrpGlyIle---ValIleIleValPheIle 477

Qy       1263 CTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCG 1322
              :::   |||   ||||||:::::::::|||:::|||   :::   |||         |||
Db        478 IleValTyrAlaAlaPheTyrAlaLeuGlyIleGlyThrValProTrp---GlnGlnSer 496

Qy       1323 GAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGC 1382
              |||:::||||||      ||||||   :::|||   |||       |||   |||
Db        497 GluLeuPheProGlnAsnValArgGlyValGlyThrSerTyrAlaThrAlaThrAsnTrp 516

Qy       1383 GTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGC 1442
                 |||   ||||||:::   ||||||||:::::   :::   ||||||
Db        517 AlaGlySerLeuValIleAlaSerThrPheLeuThrMetLeuGlnAsnIleThrProThr 536

Qy       1443 GGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTAC 1502
              |||:::|||      ::::::|||:::|||   |||:::   :::|||   |||
Db        537 GlyThrPheSerPhePheAlaGlyValAlaCysLeuSerThrIlePheCysTyrPheCys 556

Qy       1503 CTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGGCC 1562
                 ||||||      |||      |||||||||:::      :::      |||
Db        557 TyrProGluLeuSerGlyLeuGluLeuGluGluValGlnThrIleLeuLysAspGlyPhe 576

Qy       1563 GCCGCCTCGGAATCAGACGAGCCAGCCAAGGAGAAGAAGAAG 1604
                        |||         ||||||::::::|||:::
Db        577 AsnIleLysAlaSerLysAlaLeuAlaLysLysArgLysGln 590
```

```
RESULT 11
YFIG_BACSU
ID   YFIG_BACSU     STANDARD;      PRT;   482 AA.
AC   P54723;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Hypothetical metabolite transport protein yfiG.
GN   YFIG OR BSU08260.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=96262713; PubMed=8704981;
RA   Yamamoto H., Uchiyama S., Fajar A.N., Ogasawara N., Sekiguchi J.;
RT   "Determination of a 12 kb nucleotide sequence around the 76 degrees
RT   region of the Bacillus subtilis chromosome.";
RL   Microbiology 142:1417-1421(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=98044033; PubMed=9384377;
RA   Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA   Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA   Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA   Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA   Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA   Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA   Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA   Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA   Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA   Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA   Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA   Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA   Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
RA   Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA   Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA   Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA   Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA   Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA   Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA   Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA   Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA   Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA   Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA   Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA   Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA   Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
RA   Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA   Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT   "The complete genome sequence of the Gram-positive bacterium Bacillus
RT   subtilis.";
RL   Nature 390:249-256(1997).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC   -!- SIMILARITY: Belongs to the sugar transporter family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; D50543; BAA09111.1; -.
DR   EMBL; Z99108; CAB12655.1; -.
DR   PIR; B69803; B69803.
DR   SubtiList; BG11854; yfiG.
DR   InterPro; IPR007114; MFS.
DR   InterPro; IPR005828; Sub_transporter.
```

```
DR   InterPro; IPR005829; Sug_transporter.
DR   InterPro; IPR003663; Sugar_transpt.
DR   Pfam; PF00083; sugar_tr; 1.
DR   PRINTS; PR00171; SUGRTRNSPORT.
DR   TIGRFAMs; TIGR00879; SP; 1.
DR   PROSITE; PS50850; MFS; 1.
DR   PROSITE; PS00216; SUGAR_TRANSPORT_1; FALSE_NEG.
DR   PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW   Hypothetical protein; Transport; Transmembrane; Complete proteome.
FT   DOMAIN        1     29       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     30     50       1 (POTENTIAL).
FT   DOMAIN       51     59       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     60     80       2 (POTENTIAL).
FT   DOMAIN       81     92       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     93    113       3 (POTENTIAL).
FT   DOMAIN      114    120       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    121    141       4 (POTENTIAL).
FT   DOMAIN      142    155       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    156    176       5 (POTENTIAL).
FT   DOMAIN      177    184       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    185    205       6 (POTENTIAL).
FT   DOMAIN      206    263       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    264    284       7 (POTENTIAL).
FT   DOMAIN      285    301       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    302    322       8 (POTENTIAL).
FT   DOMAIN      323    331       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    332    352       9 (POTENTIAL).
FT   TRANSMEM    353    373       10 (POTENTIAL).
FT   DOMAIN      374    400       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    401    421       11 (POTENTIAL).
FT   DOMAIN      422    423       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    424    444       12 (POTENTIAL).
FT   DOMAIN      445    482       CYTOPLASMIC (POTENTIAL).
SQ   SEQUENCE   482 AA;  52756 MW;  90851C4F4C48EE01 CRC64;

Alignment Scores:
Pred. No.:              9.49e-20      Length:          482
Score:                  555.50        Matches:         152
Percent Similarity:     51.92%        Conservative:    105
Best Local Similarity:  30.71%        Mismatches:      195
Query Match:            15.37%        Indels:          44
DB:                     1             Gaps:            13

US-10-051-902A-21 (1-2017) x YFIG_BACSU (1-482)

Qy        117 GAGGCCGTCGCGCCGAAGAAG----------AAGGGCAACGTCCGGTTCGCCTTCGCC 164
              |||||||||      |||:::           ||||||   :::|||
Db          7 GluAlaValIleGlyLysGluSerLeuAlaHisLysGlyLeuLeuArg---------Thr 23

Qy        165 TGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTGATGAGC 224
                    ::::::::::::          :::|||   |||||||||   ||||||::::::
Db         24 IleThrLeuValSerThrPheGlyGlyLeuLeuPheGlyTyrAspThrGlyValIleAsn 43

Qy        225 GGGGCGTCGCTGTACATCAAGAAG------GACTTCAACATCAGTGACGGGAAGGTGGAG 278
              ||||||      ::::::                |||::::::
Db         44 GlyAlaLeuProPheMetAlaThrAlaGlyGlnLeuAsnLeuThrProValThrGluGly 63

Qy        279 GTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACG 338
              ::::::          |||   |||   :::      |||:::          ||||||
Db         64 LeuValAlaSerSerLeuLeuLeuGlyAlaAlaPheGlyAlaMetPheGlyGlyArgLeu 83

Qy        339 TCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGG 398
              ||||||      |||||||||   ||||||::::::   |||::::::|||   |||
Db         84 SerAspArgHisGlyArgArgLysThrIleLeuTyrLeuAlaLeuLeuPheIleAlaAla 103

Qy        399 GSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCC 458
                          |||:::   |||   ::::::::::::      ||||||:::
Db        104 ThrLeuGlyCysThrPheSerProAsnAlaSerValMetIleAlaPheArgPheLeuLeu 123

Qy        459 GGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCG 518
              |||:::   ||||||   |||   :::      |||   :::   ||||||:::||||||
Db        124 GlyLeuAlaValGlyCysAlaSerValThrValProThrPheLeuAlaGluIleSerPro 143

Qy        519 GCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTG 578
              |||      ||||||   :::   :::      |||:::   |||   .   |||   |||
Db        144 AlaGluArgArgGlyArgIleValThrGlnAsnGluLeuMetIleValIleGlyGlnLeu 163

Qy        579 CTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGG--------- 629
              |||   |||      |||          :::   :::          :::|||
Db        164 LeuAlaTyrThrPheAsn---------AlaIleIleGlySerThrMetGlyGluSerAla 180

Qy        630 ------TGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATG 683
                    ||||||   ||||||   |||          |||:::|||:::|||
Db        181 AsnValTrpArgTyrMetLeuValIleAlaThrLeuProAlaValValLeuTrpPheGly 200

Qy        684 GTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCC 743
              :::|||   :::|||||||||||||||||||||||      |||||||||:::   ||||||
Db        201 MetLeuIleValProGluSerProArgTrpLeuAlaAlaLysGlyArgMetGlyAspAla 220

Qy        744 AAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGAC 803
                    ||||||   :::      :::   :::   :::|||   :::   :::   :::
Db        221 LeuArgValLeuArgGlnIleArgGluAspSer---GlnAlaGlnGlnGluIleLysGlu 239

Qy        804 ATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAG 863
              ||||||   |||                  ::::::|||          |||      |||
Db        240 IleLysHisAla-----------------IleGluGly---------ThrAlaLysLys 250

Qy        864 AGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCC 923
                  |||      :::   :::      |||
Db        251 AlaGlyPheHisAspPheGlnGluProTrp---------------------------- 260

Qy        924 ATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATT 983
              :::||||||||||      ||||||||||      |||      |||:::
Db        261 IleArgArgIleLeuPheIleGlyIleGlyIleAlaIleValGlnGlnIleThrGlyVal 280

Qy        984 CACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAA 1043
              :::|||::::::::::|||          :::   :::      |||      :::
Db        281 AsnSerIleMetTyrTyrGlyThrGluIleLeuArgGluAlaGlyPheGlnThrGluAla 300

Qy       1044 CACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCG 1103
                    :::|||          ||||||          :::   ::::::
Db        301 AlaLeuIleGlyAsnIle---AlaAsnGlyValIleSerValIleAlaValIlePheGly 319

Qy       1104 ACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATA 1163
                 :::   :::      |||   |||||||||:::|||:::          |||
Db        320 IleTrpLeuLeuGlyLysValArgArgArgProMetLeuIleIleGlyGlnIleGlyThr 339

Qy       1164 ATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCC 1223
                 :::   :::|||:::   :::|||      |||:::::::|||   :::         |||
Db        340 MetThrAlaLeuLeuLeuIleGlyIle---LeuSerIleValLeuGluGlyThrProAla 358

Qy       1224 AAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTC 1283
                 :::|||:::   :::          ::::::   |||:::   ::::::||||||
Db        359 ---LeuProTyrValVal---------LeuSerLeuThrIleLeuPheLeuAlaPheGln 374

Qy       1284 TCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTG 1343
                    :::      :::||||||:::      ||||||||||||||:::   |||
Db        375 GlnThrAlaIleSerThrValThrTrpLeuMetLeuSerGluIlePheProMetHisVal 394

Qy       1344 CGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCC 1403
              |||   ||||||      :::                  :::   :::|||
Db        395 ArgGlyLeuGlyMetGlyIleSerThrPheCysLeuTrpThrAlaAsnPheLeuIleGly 414

Qy       1404 ATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCC 1463
                 ||||||      |||          |||   :::      :::||||||::::::
Db        415 PheThrPheProIleLeuLeuAsnHisIleGlyMetSerAlaThrPhePheIlePheVal 434

Qy       1464 GGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGG 1523
                 :::      ||||||   :::|||          |||:::||||||||||:::||||||
Db        435 AlaMetAsnIleLeuAlaIleLeuPheValLysLysTyrValProGluThrLysGlyArg 454

Qy       1524 ACGCTGGAGGAGAT-GAGCAAGCTGTTCGGCGACACGGCCGCCGC 1567
                 :::|||||:::   |||   :::   |||::::::||||||||
Db        455 SerLeuGluGlnLeuGluHisSerPheArgGlnTyrGlyArgArg 469

RESULT 12
```

```
STP1_ARATH
ID   STP1_ARATH     STANDARD;      PRT;   522 AA.
AC   P23586; Q9SXB1;
DT   01-NOV-1991 (Rel. 20, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Glucose transporter (Sugar carrier).
GN   STP1 OR AT1G11260 OR T28P6.9.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Landsberg erecta;
RX   MEDLINE=91005995; PubMed=2209537;
RA   Sauer N., Friedlaender K., Graeml-Wicke U.;
RT   "Primary structure, genomic organization and heterologous expression
RT   of a glucose transporter from Arabidopsis thaliana.";
RL   EMBO J. 9:3045-3050(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=21016719; PubMed=11130712;
RA   Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA   White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA   Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA   Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA   Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA   Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA   Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA   Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA   Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA   Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziali A.,
RA   Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA   Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA   Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA   Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA   Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA   Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT   "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 408:816-820(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=22954850; PubMed=14593172;
RA   Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,
RA   Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,
RA   Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
RA   Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
RA   Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,
RA   Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,
RA   Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,
RA   Hayashizaki Y., Johnson-Hopson C., Hsuan V.W., Iida K., Karnes M.,
RA   Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,
RA   Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
RA   Satou M., Tamse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,
RA   Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT   "Empirical analysis of transcriptional activity in the Arabidopsis
RT   genome.";
RL   Science 302:842-846(2003).
CC   -!- FUNCTION: Active uptake of hexoses. Probable glucose/hydrogen
CC       symport.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to the sugar transporter family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X55350; CAA39037.1; -.
DR   EMBL; AC007259; AAD49995.1; -.
DR   EMBL; AY059781; AAL24129.1; -.
DR   EMBL; AY054249; AAL06908.1; -.
DR   PIR; E86246; E86246.
DR   PIR; S12042; S12042.
DR   InterPro; IPR007114; MFS.
DR   InterPro; IPR005828; Sub_transporter.
DR   InterPro; IPR005829; Sug_transporter.
DR   InterPro; IPR003663; Sugar_transpt.
DR   Pfam; PF00083; sugar_tr; 1.
DR   PRINTS; PR00171; SUGRTRNSPORT.
DR   TIGRFAMs; TIGR00879; SP; 1.
DR   PROSITE; PS50850; MFS; 1.
DR   PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR   PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW   Transmembrane; Transport; Sugar transport; Symport.
FT   DOMAIN        1     23       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     24     44       1 (POTENTIAL).
FT   TRANSMEM     84    104       2 (POTENTIAL).
FT   TRANSMEM    118    138       3 (POTENTIAL).
FT   TRANSMEM    141    161       4 (POTENTIAL).
FT   TRANSMEM    170    190       5 (POTENTIAL).
FT   TRANSMEM    203    223       6 (POTENTIAL).
FT   TRANSMEM    296    318       7 (POTENTIAL).
FT   TRANSMEM    325    345       8 (POTENTIAL).
FT   TRANSMEM    349    369       9 (POTENTIAL).
FT   TRANSMEM    385    405       10 (POTENTIAL).
FT   TRANSMEM    431    451       11 (POTENTIAL).
FT   TRANSMEM    454    474       12 (POTENTIAL).
FT   DOMAIN      475    522       CYTOPLASMIC (POTENTIAL).
FT   CONFLICT    333    333       A -> G (IN REF. 1).
SQ   SEQUENCE   522 AA;  57610 MW;  813AFC8AFA7AA583 CRC64;

Alignment Scores:
Pred. No.:                3.57e-19        Length:          522
Score:                    543.50          Matches:         146
Percent Similarity:       48.25%          Conservative:    102
Best Local Similarity:    28.40%          Mismatches:      206
Query Match:              15.04%          Indels:          60
DB:                       1               Gaps:            13

US-10-051-902A-21 (1-2017) x STP1_ARATH (1-522)

Qy          92 GGAGATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCG 151
               |||||||||       |||   |||     :::   |||
Db           9 GlyAspGly-GlnLysAlaTyrProGlyLysLeuThrPro---------------------- 21

Qy         152 GTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATAT 211
               |||           :::::|||:::|||     ::::::   ||||||||||
Db          22 -PheValLeuPheThrCysValValAlaAlaMetGlyGlyLeuIlePheGlyTyrAspIl 41

Qy         212 CGGGGTGATGAGCGGGGCG----------------------------------TCGCT 235
               ||||:::     |||                                      |||::
Db          41 eGlyIleSerGlyGlyValThrSerMetProSerPheLeuLysArgPhePheProServa 61

Qy         236 GTACATCAAGAAGGACTTCAACATCAGTGACGGGAAG---------------------- 272
               :|||   |||:::     :::   |||     :::
Db          61 lTyrArgLysGlnGlnGluAspAlaSerThrAsnGlnTyrCysGlnTyrAspSerProTh 81

Qy         273 -GTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGG 331
                :::   :::           |||   |||   :::||||||   |||     |||
Db          81 rLeuThrMetPheThrSerSerLeuTyrLeuAlaAlaLeuIleSerSerLeuValAlaSe 101

Qy         332 GCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTT 391
                   :::       |||||||||   :::::::::|||     ::::::|||
Db         101 rThrValThrArgLysPheGlyArgArgLeuSerMetLeuPheGlyGlyIleLeuPheCy 121

Qy         392 CGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTT 451
               ||||||      :::   |||||||||   :::     ||||||:::   ||||||
```

```
Db        121 sAlaGlyAlaLeuIleAsnGlyPheAlaLysHisValTrpMetLeuIleValGlyArgIl 141

Qy        452 CGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGT 511
              :::   |||   |||:::|||:::|||             |||:::|||   :::|||::
Db        141 eLeuLeuGlyPheGlyIleGlyPheAlaAsnGlnAlaValProLeuTyrLeuSerGluMe 161

Qy        512 GTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGG 571
              ::::|||         ||||||   |||             ::::::   |||      ||
Db        161 tAlaProTyrLysTyrArgGlyAlaLeuAsnIleGlyPheGlnLeuSerIleThrIleGl 181

Qy        572 CATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTG 631
              |||||||:::      |||   ||||||   |||:::::::::             |||||
Db        181 yIleLeuValAlaGluValLeuAsnTyrPhePheAlaLysIleLysGlyGlyTrpGlyTr 201

Qy        632 GCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGG 691
              ||||:::   ||||||           |||:::::::::::   :::      |||
Db        201 pArgLeuSerLeuGlyGlyAlaValValProAlaLeuIleIleThrIleGlySerLeuVa 221

Qy        692 CATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGT 751
              :::|||:::::|||      ::::::   :::|||:::
Db        221 lLeuProAspThrProAsnSerMetIleGluArgGlyGln-------------------- 234

Qy        752 GCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGC 811
                                          |||||||||      :::|||      |||:::
Db        235 ----------------------HisGluGluAlaLysThrLysLeuArgArgIleArgGl 247

Qy        812 CGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAG 871
                    :::   :::|||   |||   |||:::|||         |||
Db        247 yValAspAspValSerGlnGluPheAsp---AspLeuValAlaAlaSerLysGluSerGl 266

Qy        872 CGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCG 931
                  |||         |||:::   |||:::                        |||
Db        266 nSerIleGlu---HisProTrpArgAsnLeuLeu---------ArgArgLysTyrArgPr 282

Qy        932 CATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGT 991
                |||         :::   |||   |||||||||         ||||||:::   ::
Db        282 oHisLeuThrMetAlaValMetIleProPhePheGlnGlnLeuThrGlyIleAsnValIl 302

Qy        992 CGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTT 1051
              ::::||||||:::|||:::::|||   :::   |||   |||   |||         ::
Db        302 eMetPheTyrAlaProValLeuPheAsnThrIleGlyPheThrThrAspAlaSerLeuMe 322

Qy       1052 GGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATC---------TTGTTGGC 1102
              :                  ||||||      :::   :::          |||:::::
Db        322 tSerAlaVal------------ValThrGlySerValAsnValAlaAlaThrLeuValSe 338

Qy       1103 GACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTG---GGCAGCACGGGCGG 1159
              :   :::   :::|||      |||||||||   |||   |||   |||   |||
Db        338 rIleTyrGlyValAspArgTrpGlyArgArgPheLeuPheLeuGluGlyGlyThrGlnMe 358

Qy       1160 GATAATCCTCTCCCTCATC---------GGCCTCGGCGCCGGGCTCACCGTCGTCGGCCA 1210
               :::|||          :::             :::||||||          |||   |||
Db        358 tLeuIleCysGlnAlaValValAlaAlaCysIleGlyAlaLysPheGlyValAspGlyTh 378

Qy       1211 GCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTA 1270
              |||                 |||         :::   :::                ||
Db        378 r---ProGlyGluLeuProLysTrpTyrAlaIleValValValThrPheIleCysIleTy 397

Qy       1271 CGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTT 1330
              |||||||   |||:::         ||||||:::   |||:::      ||||||||||
Db        397 rValAlaGlyPheAlaTrpSerTrpGlyProLeuGlyTrpLeuValProSerGluIlePh 417

Qy       1331 CCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAG 1390
              |||||||:::::|||:::         |||:::   |||:::   |||   :::   ::
Db        417 eProLeuGluIleArgSerAlaAlaGlnSerIleThrValSerValAsnMetIlePheTh 437

Qy       1391 CGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTT 1450
              :   :::|||:::      ||||||::::::         :::         |||   ||
Db        437 rPheIleIleAlaGlnIlePheLeuThrMetLeuCysHisLeuLysPhe---GlyLeuPh 456

Qy       1451 CTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGA 1510
              |   :::::::::             :::      :::|||   |||   :::|||||||||
Db        456 eLeuValPheAlaPhePheValValValMetSerIlePheValTyrIlePheLeuProGl 476

Qy       1511 GACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTC 1550
                  ||||:::|||      :::||||||||||   :::::::::
Db        476 uThrLysGlyIleProIleGluGluMetGlyGlnValTrp 489

RESULT 13
HUP3_CHLKE
ID   HUP3_CHLKE     STANDARD;      PRT;   534 AA.
AC   Q39525;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   H(+)/hexose cotransporter 3.
GN   HUP3.
OS   Chlorella kessleri.
OC   Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;
OC   Chlorellales incertae sedis; Parachlorella.
OX   NCBI_TaxID=3074;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95175627; PubMed=7870840;
RA   Stadler R., Wolf K., Hilgarth C., Tanner W., Sauer N.K.;
RT   "Subcellular localization of the inducible Chlorella HUP1
RT   monosaccharide-H+ symporter and cloning of a Co-induced galactose-H+
RT   symporter.";
RL   Plant Physiol. 107:33-41(1995).
CC   -!- FUNCTION: Active uptake of hexoses.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to the sugar transporter family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X75440; CAA53192.1; -.
DR   PIR; S38435; S38435.
DR   InterPro; IPR007114; MFS.
DR   InterPro; IPR005828; Sub_transporter.
DR   InterPro; IPR005829; Sug_transporter.
DR   InterPro; IPR003663; Sugar_transpt.
DR   Pfam; PF00083; sugar_tr; 1.
DR   PRINTS; PR00171; SUGRTRNSPORT.
DR   TIGRFAMs; TIGR00879; SP; 1.
DR   PROSITE; PS50850; MFS; 1.
DR   PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR   PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW   Transmembrane; Transport; Sugar transport; Symport; Repeat;
KW   Multigene family.
FT   DOMAIN        1     20       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     21     41       1 (POTENTIAL).
FT   DOMAIN       42     86       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     87    107       2 (POTENTIAL).
FT   DOMAIN      108    113       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    114    135       3 (POTENTIAL).
FT   DOMAIN      136    137       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    138    158       4 (POTENTIAL).
FT   DOMAIN      159    173       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    174    194       5 (POTENTIAL).
FT   DOMAIN      195    204       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    205    225       6 (POTENTIAL).
FT   DOMAIN      226    299       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    300    322       7 (POTENTIAL).
FT   DOMAIN      323    328       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    329    349       8 (POTENTIAL).
FT   DOMAIN      350    352       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    353    373       9 (POTENTIAL).
FT   DOMAIN      374    394       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    395    415       10 (POTENTIAL).
FT   DOMAIN      416    433       CYTOPLASMIC (POTENTIAL).
```

```
FT   TRANSMEM     434     454       11 (POTENTIAL).
FT   DOMAIN       455     456       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     457     477       12 (POTENTIAL).
FT   DOMAIN       478     534       CYTOPLASMIC (POTENTIAL).
SQ   SEQUENCE   534 AA;  57772 MW;  52E03715F3726017 CRC64;

Alignment Scores:
Pred. No.:                5.86e-19        Length:          534
Score:                    539.00          Matches:         155
Percent Similarity:       47.70%          Conservative:    83
Best Local Similarity:    31.06%          Mismatches:      205
Query Match:              14.91%          Indels:          56
DB:                       1               Gaps:            16

US-10-051-902A-21 (1-2017) x HUP3_CHLKE (1-534)

Qy        141 GGCAACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTC 200
Db         21 GlyLeuThrAlaTyrValLeuLeuValAlaLeuValAlaAlaCysGlyGlyMetLeuLeu  40

Qy        201 GGCTACGATATCGGGGTGATGAGCGGGGCGTCG-------------------------- 233
Db         41 GlyTyrAspAsnGlyValThrGlyGlyValAlaSerMetGluGlnPheGluArgLysPhe  60

Qy        234 ---------CTGTACATCAAGAAGGAC----------------------TTCAACATC 260
Db         61 PheProAspValTyrGluLysLysGlnGlnIleValGluThrSerProTyrCysThrTyr  80

Qy        261 AGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCC 320
Db         81 AspAsnProLysLeuGlnLeuPheValSerSerLeuPheLeuAlaGlyLeuIleSerCys 100

Qy        321 ---TTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCC 377
Db        101 IlePheSerAlaTrpIleThrArgAsnTrp---GlyArgLysAlaSerMetGlyIleGly 119

Qy        378 GCCGTCATATTCTTC---GCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATG 434
Db        120 GlyIlePhePheIleAlaAlaGlyGlyLeuValAsnAlaPheAlaGlnAspIleAlaMet 139

Qy        435 CTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCG 494
Db        140 LeuIleValGlyArgValLeuLeuGlyPheGlyValGlyLeuGlySerGlnValValPro 159

Qy        495 GTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAG 554
Db        160 GlnTyrLeuSerGluValAlaProPheSerHisArgGlyMetLeuAsnIleGlyTyrGln 179

Qy        555 GTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTG 614
Db        180 LeuPheValThrIleGlyIleLeuIleAlaGlyLeuValAsnTyrGlyValArgAsn--- 198

Qy        615 CCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTC 674
Db        199 ---TrpAspAsnGlyTrpArgLeuSerLeuGlyLeuAlaAlaValProGlyLeuIleLeu 217

Qy        675 GCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTC 734
Db        218 LeuLeuGlyAlaIleValLeuProGluSerProAsnPheLeuValGluLysGlyArgThr 237

Qy        735 GCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGC 794
Db        238 AspGlnGlyArgArgIleLeuGluLysLeuArgGlyThrSerHisValGluAlaGlu--- 256

Qy        795 CTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACC 854
Db        257 PheAlaAspIleValAlaAlaValGluIleAlaArgProIle------------------ 270

Qy        855 GTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCG 914
Db        271 ---------------------ThrMetArgGlnSerTrpArgSerLeuPhe--------- 280

Qy        915 ACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCG 974

Db        281 ThrArgArgTyrMetProGlnLeuLeuThrSerPheValIleGlnPhePheGlnGlnPhe 300

Qy        975 TTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACG 1034
Db        301 ThrGlyIleAsnAlaIleIlePheTyrValProValLeuPheSerSerLeuGlySerAla 320

Qy       1035 AACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATC 1094
Db        321 SerSerAlaAlaLeuLeuAsnThrVal---ValValGlyAlaValAsnValGlySerThr 339

Qy       1095 TTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACG 1154
Db        340 MetIleAlaValLeuLeuSerAspLysPheGlyArgArgPheLeuLeuIle------Glu 357

Qy       1155 GGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTC---------GTC 1205
Db        358 GlyGlyIleThrCysCysLeuAlaMetLeuAlaAlaGlyIleThrLeuGlyValGluPhe 377

Qy       1206 GGCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTC 1265
Db        378 GlyGlnTyrGlyThrGluAspLeuProHisProValSerAlaGlyValLeuAlaValIle 397

Qy       1266 GCC---TACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCG 1322
Db        398 CysIlePheIleAlaGlyPheAlaTrpSerTrpGlyProMetGlyTrpLeuIleProSer 417

Qy       1323 GAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGC 1382
Db        418 GluIlePheThrLeuGluThrArgProAlaGlyThrAlaValAlaValMetGlyAsnPhe 437

Qy       1383 GTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGC 1442
Db        438 LeuPheSerPheValIleGlyGlnAlaPheValSerMetLeuCysAlaMetLysPhe--- 456

Qy       1443 GGCAGCTTCTTCCTCTACTCCGGC---------ATCGCCGCGCTCGCCTGGGTGTTCTTC 1493
Db        457 GlyValPheLeuPhePheAlaGlyTrpLeuValIleMetValLeuCysAlaIlePheLeu 476

Qy       1494 TACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTC 1550
Db        477 ---------LeuProGluThrLysGlyValProIleGluArgValGlnAlaLeuTyr 492

RESULT 14
HUP1_CHLKE
ID   HUP1_CHLKE     STANDARD;      PRT;   534 AA.
AC   P15686; Q39523;
DT   01-APR-1990 (Rel. 14, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   H(+)/hexose cotransporter 1.
GN   HUP1.
OS   Chlorella kessleri.
OC   Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;
OC   Chlorellales incertae sedis; Parachlorella.
OX   NCBI_TaxID=3074;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90092536; PubMed=2599110;
RA   Sauer N., Tanner W.;
RT   "The hexose carrier from Chlorella. cDNA cloning of a eucaryotic H+-
RT   cotransporter.";
RL   FEBS Lett. 259:43-46(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91330335; PubMed=1868571;
RA   Wolf K., Tanner W., Sauer N.;
RT   "The Chlorella H+/hexose cotransporter gene.";
RL   Curr. Genet. 19:215-219(1991).
CC   -!- FUNCTION: Active uptake of hexoses.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- INDUCTION: By glucose.
CC   -!- SIMILARITY: Belongs to the sugar transporter family.
```

```
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; Y07520; CAA68813.1; -.
DR   EMBL; X55349; CAA39036.1; -.
DR   PIR; S14144; S14144.
DR   InterPro; IPR007114; MFS.
DR   InterPro; IPR005828; Sub_transporter.
DR   InterPro; IPR005829; Sug_transporter.
DR   InterPro; IPR003663; Sugar_transpt.
DR   Pfam; PF00083; sugar_tr; 1.
DR   PRINTS; PR00171; SUGRTRNSPORT.
DR   TIGRFAMs; TIGR00879; SP; 1.
DR   PROSITE; PS50850; MFS; 1.
DR   PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR   PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW   Transmembrane; Transport; Sugar transport; Symport; Repeat;
KW   Multigene family.
FT   DOMAIN        1     21       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     22     42       1 (POTENTIAL).
FT   DOMAIN       43     87       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     88    108       2 (POTENTIAL).
FT   DOMAIN      109    114       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    115    135       3 (POTENTIAL).
FT   DOMAIN      136    144       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    145    165       4 (POTENTIAL).
FT   DOMAIN      166    173       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    174    194       5 (POTENTIAL).
FT   DOMAIN      195    204       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    205    225       6 (POTENTIAL).
FT   DOMAIN      226    299       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    300    322       7 (POTENTIAL).
FT   DOMAIN      323    328       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    329    349       8 (POTENTIAL).
FT   DOMAIN      350    352       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    353    373       9 (POTENTIAL).
FT   DOMAIN      374    387       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    388    408       10 (POTENTIAL).
FT   DOMAIN      409    433       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    434    454       11 (POTENTIAL).
FT   DOMAIN      455    456       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    457    477       12 (POTENTIAL).
FT   DOMAIN      478    534       CYTOPLASMIC (POTENTIAL).
FT   CONFLICT     60     60       MISSING (IN REF. 1).
FT   CONFLICT    209    209       P -> L (IN REF. 1).
FT   CONFLICT    248    248       C -> R (IN REF. 1).
SQ   SEQUENCE   534 AA;  57522 MW;  3B238E8710AA2237 CRC64;

Alignment Scores:
Pred. No.:                6.19e-19      Length:          534
Score:                    538.50        Matches:         166
Percent Similarity:       46.48%        Conservative:    78
Best Local Similarity:    31.62%        Mismatches:      206
Query Match:              14.90%        Indels:          75
DB:                       1             Gaps:            18

US-10-051-902A-21 (1-2017) x HUP1_CHLKE (1-534)

Qy        168 GCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGG 227
              |||   :::|||:::         :::|||||||||||||||   ||||||     |||
Db         31 AlaPheMetAlaAlaCysGlyGlyLeuLeuLeuGlyTyrAspAsnGlyValThrGlyGly 50

Qy        228 GCG---------TCGCTGTACATCAAGAAGGACTTC------------------------ 254
                          :::   ||||||   |||
Db         51 ValValSerLeuGluAlaPheGluLysLysPhePheProAspValTrpAlaLysLysGln 70

Qy        255 --------------------------AACATCAGTGACGGGAAGGTGGAGGTTCTCATG 287
```

```
                                                  :::   |||:::::::::   :::
Db         71 GluValHisGluAspSerProTyrCysThrTyrAspAsnAlaLysLeuGlnLeuPheVal 90

Qy        288 GGCATACTGAACCTCTACTCGCTCATCGGCTCC---TTCGCGGCGGGGCGGACGTCGGAC 344
                 |||   |||      |||:::        ||||||:::     |||   :::
Db         91 SerSerLeuPheLeuAlaGlyLeuValSerCysLeuPheAlaSerTrpIleThrArgAsn 110

Qy        345 TGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGSGTTC 404
              |||   ||||||:::   |||:::                    |||   ||||||
Db        111 Trp---GlyArgLysValThrMetGlyIleGlyGlyAlaPhePheValAlaGlyGlyLeu 129

Qy        405 CTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATC 464
              :::     ||||||   :::   |||||||||:::   ||||||   :::   |||
Db        130 ValAsnAlaPheAlaGlnAspMetAlaMetLeuIleValGlyArgValLeuGlyPhe 149

Qy        465 GGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCG 524
              |||||||||           :::   |||   |||   :::||||||:::|||   |||
Db        150 GlyValGlyLeuGlySerGlnValValProGlnTyrLeuSerGluValAlaProPheSer 169

Qy        525 GCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGG 584
                 ||||||   |||           ::::::|||:::     |||||||||:::
Db        170 HisArgGlyMetLeuAsnIleGlyTyrGlnLeuPheValThrIleGlyIleLeuIleAla 189

Qy        585 TACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTC 644
                :::   ||||||||                 |||||||||:::   |||
Db        190 GlyLeuValAsnTyrAlaValArg------AspTrpGluAsnGlyTrpArgLeuSerLeu 207

Qy        645 GGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCG 704
              |||      |||||||||      :::|||   |||      |||   :::||||||||||
Db        208 GlyProAlaAlaAlaProGlyAlaIleLeuPheLeuGlySerLeuValLeuProGluSer 227

Qy        705 CCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACC 764
              |||   :::||||||   ||||||:::          :::   ||||||:::|||
Db        228 ProAsnPheLeuValGluLysGlyLysThrGluLysGlyArgGluValLeuGlnLysLeu 247

Qy        765 TCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATC 824
                  |||:::|||      ||||||      |||||||||   ||||||      |||
Db        248 CysGlyThrSerGluValAspAlaGlu---PheAlaAspIleValAlaAlaValGluIle 266

Qy        825 CCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAG 884
                    :::                                              :::
Db        267 AlaArgProIle-------------------------------------------ThrMetArg 273

Qy        885 CGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCC 944
              :::   |||      |||            |||                  ||||||:::
Db        274 GlnSerTrpAlaSerLeuPhe---------ThrArgArgTyrMetProGlnLeuLeuThr 290

Qy        945 GGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGC 1004
                  |||   |||||||||         ||||||:::::::::::::||||||
Db        291 SerPheValIleGlnPhePheGlnGlnPheThrGlyIleAsnAlaIleIlePheTyrVal 310

Qy       1005 CCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGCACCACTTGG 1064
              |||::::::|||   |||   |||      |||           |||   |||
Db        311 ProValLeuPheSerSerLeuGlySerAlaAsnSerAlaAlaLeuLeuAsnThrVal--- 329

Qy       1065 CCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTC 1124
                 |||                  |||:::|||      |||   |||
Db        330 ValValGlyAlaValAsnValGlySerThrLeuIleAlaValMetPheSerAspLysPhe 349

Qy       1125 GGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTC 1184
              |||||||||   ||||||:::          |||||||||           |||   |||
Db        350 GlyArgArgPheLeuLeuIle------GluGlyGlyIleGlnCysCysLeuAlaMetLeu 367

Qy       1185 GGCGCCGGGCTCACCGTCGTC-----------------GGCCAGCACCCCGACGCCAAG 1226
                 |||:::   :::                      |||      |||
Db        368 ThrThrGlyValValLeuAlaIleGluPheAlaLysTyrGlyThrAspPro--------- 384

Qy       1227 ATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCC---TACGTCGCCTTCTTC 1283
              :::|||   |||:::        |||   :::   :::      :::::::::   |||
Db        385 LeuProLysAlaValAlaSerGlyIleLeuAlaValIleCysIlePheIleSerGlyPhe 404

Qy       1284 TCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTG 1343
              :::         ||||||:::   |||:::      ||||||||||||   |||:::
```

```
Db        405 AlaTrpSerTrpGlyProMetGlyTrpLeuIleProSerGluIlePheThrLeuGluThr 424

Qy       1344 CGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCC 1403
              |||       |||   ::::::   |||       |||   :::   |||   ||||||
Db        425 ArgProAlaGlyThrAlaValAlaValValGlyAsnPheLeuPheSerPheValIleGly 444

Qy       1404 ATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCC 1463
                 |||:::|||:::      |||:::         |||   |||      ::::::
Db        445 GlnAlaPheValSerMetLeuCysAlaMetGluTyr---GlyValPheLeuPhePheAla 463

Qy       1464 GGC---------ATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACC 1514
              |||         |||      |||      :::|||              ||||||||||||
Db        464 GlyTrpLeuValIleMetValLeuCysAlaIlePheLeu---------LeuProGluThr 480

Qy       1515 CGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTC---------------------- 1550
              :::|||     :::|||   :::      |||:::
Db        481 LysGlyValProIleGluArgValGlnAlaLeuTyrAlaArgHisTrpPheTrpAsnArg 500

Qy       1551 ------GGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAGCCAAGGAGAAGAAGAAG 1604
                 |||      ||||||          :::      |||:::::::::|||:::
Db        501 ValMetGlyProAlaAlaAla-----------GluValIleAlaGluAspGluLysArg 516

Qy       1605 GTGGAAATGGCCGCC 1619
              |||      |||:::
Db        517 ValAlaAlaAlaSer 521

RESULT 15
GLCP_SYNY3
ID   GLCP_SYNY3     STANDARD;      PRT;   468 AA.
AC   P15729;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Glucose transport protein.
GN   GTR OR GLCP OR SLL0771.
OS   Synechocystis sp. (strain PCC 6803).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX   NCBI_TaxID=1148;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90014182; PubMed=2507869;
RA   Zhang C.C., Durand M.C., Jeanjean R., Joset F.;
RT   "Molecular and genetical analysis of the fructose-glucose transport
RT   system in the cyanobacterium Synechocystis PCC6803.";
RL   Mol. Microbiol. 3:1221-1229(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91346660; PubMed=2129397;
RA   Schmetterer G.R.;
RT   "Sequence conservation among the glucose transporter from the
RT   cyanobacterium Synechocystis sp. PCC 6803 and mammalian glucose
RT   transporters.";
RL   Plant Mol. Biol. 14:697-706(1990).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96127529; PubMed=8590279;
RA   Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA   Sugiura M., Tabata S.;
RT   "Sequence analysis of the genome of the unicellular cyanobacterium
RT   Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT   region from map positions 64% to 92% of the genome.";
RL   DNA Res. 2:153-166(1995).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to the sugar transporter family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X15988; CAA34119.1; -.
DR   EMBL; X16472; CAA34492.1; -.
DR   EMBL; D64000; BAA10117.1; -.
DR   PIR; S10014; S10014.
DR   InterPro; IPR007114; MFS.
DR   InterPro; IPR005828; Sub_transporter.
DR   InterPro; IPR005829; Sug_transporter.
DR   InterPro; IPR003663; Sugar_transpt.
DR   Pfam; PF00083; sugar_tr; 1.
DR   PRINTS; PR00171; SUGRTRNSPORT.
DR   TIGRFAMs; TIGR00879; SP; 1.
DR   PROSITE; PS50850; MFS; 1.
DR   PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR   PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW   Transmembrane; Transport; Sugar transport; Complete proteome.
FT   DOMAIN        1     17      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     18     38      1 (POTENTIAL).
FT   DOMAIN       39     58      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     59     78      2 (POTENTIAL).
FT   DOMAIN       79     84      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     85    105      3 (POTENTIAL).
FT   DOMAIN      106    114      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    115    135      4 (POTENTIAL).
FT   DOMAIN      136    149      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    150    170      5 (POTENTIAL).
FT   DOMAIN      171    186      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    187    207      6 (POTENTIAL).
FT   DOMAIN      208    265      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    266    286      7 (POTENTIAL).
FT   DOMAIN      287    307      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    308    328      8 (POTENTIAL).
FT   DOMAIN      329    331      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    332    352      9 (POTENTIAL).
FT   DOMAIN      353    366      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    367    387      10 (POTENTIAL).
FT   DOMAIN      388    412      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    413    433      11 (POTENTIAL).
FT   DOMAIN      434    436      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    437    457      12 (POTENTIAL).
FT   DOMAIN      458    468      CYTOPLASMIC (POTENTIAL).
FT   CONFLICT    319    319      L -> I (IN REF. 1).
SQ   SEQUENCE   468 AA;  49747 MW;  D7EC545C4FB38D22 CRC64;

Alignment Scores:
Pred. No.:                1.33e-18        Length:          468
Score:                    531.50          Matches:         149
Percent Similarity:       48.58%          Conservative:    90
Best Local Similarity:    30.28%          Mismatches:      210
Query Match:              14.71%          Indels:          43
DB:                       1               Gaps:            11

US-10-051-902A-21 (1-2017) x GLCP_SYNY3 (1-468)

Qy        114 CCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGCGCCATC 173
              |||   :::         :::         ||||||:::|||           :::
Db          3 ProSerSerSerProSerGlnSerThrAlaAsnValLysPheValLeuLeuIleSerGly 22

Qy        174 CTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCG 233
              :::|||:::::         |||   |||:::|||      |||::::::||||||
Db         23 ValAlaAlaLeuGlyGlyPheLeuPheGlyPheAspThrAlaValIleAsnGlyAlaVal 42

Qy        234 CTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATA 293
              ::::::|||   |||                        :::   :::.  :::
Db         43 AlaAlaLeuGlnLysHisPheGlnThrAspSerLeuLeuThrGlyLeuSerValSerLeu 62

Qy        294 CTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGC 353
              |||   |||   :::|||:::|||   ||||||      :::|||      |||
Db         63 AlaLeuLeuGlySerAlaLeuGlyAlaPheGlyAlaGlyProIleAlaAspArgHisGly 82

Qy        354 CGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGGSGTTCCTCATGGGG 413
              |||      |||::::::   |||||||||:::|||                   |||
Db         83 ArgIleLysThrMetIleLeuAlaAlaValLeuPheThrLeuSerSerIleGlySerGly 102
```

```
Qy       414 TTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGC 473
                                 :::|||  |||  :::  ||||||||||||||||
Db       103 LeuProPheThrIleTrpAspPheIlePheTrpArgValLeuGlyGlyIleGlyValGly 122

Qy       474 TACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGC 533
               |||   :::||||||||||   |||   ||||||||||||||||||      ||||||
Db       123 AlaAlaSerValIleAlaProAlaTyrIleAlaGluValSerProAlaHisLeuArgGly 142

Qy       534 TTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCG 593
               |||   |||      ::::::   |||      ||||||   :::      :::|||
Db       143 ArgLeuGlySerLeuGlnGlnLeuAlaIleValSerGlyIlePheIleAlaLeuLeuSer 162

Qy       594 AACTATGCTTTCTCCCGCTTG--------------------CCGCTGAACCTCGGG--- 629
             |||:::      :::   :::                   |||        |||
Db       163 AsnTrpPheIleAlaLeuMetAlaGlyGlySerAlaGlnAsnProTrpLeuPheGlyAla 182

Qy       630 ------TGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATG 683
                   ||||||   |||                   |||::::::|||      :::
Db       183 AlaAlaTrpArgTrpMetPheTrpThrGluLeuIleProAlaLeuLeuTyrGlyValCys 202

Qy       684 GTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCC 743
                 :::||||||||||||||||:::||||||   :::|||:::          |||
Db       203 AlaPheLeuIleProGluSerProArgTyrLeuValAlaGlnGlyGlnGlyGluLysAla 222

Qy       744 AAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGAC 803
               :::|||   |||                 :::         |||:::   :::
Db       223 AlaAlaIleLeuTrpLysValGlu------GlyGlyAspValProSerArgIleGluGlu 240

Qy       804 ATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAG 863
             |||:::|||          :::      :::                    |||:::
Db       241 IleGlnAlaThrValSerLeuAspHisLys----------------------ProArg 252

Qy       864 AGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCC 923
                                           :::||||||
Db       253 PheSerAsp----------------------------LeuLeuSerArgArgGlyGly 262

Qy       924 ATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATT 983
             :::      |||:::      |||:::|||:::          |||      :::||||||
Db       263 LeuLeuProIleValTrpIleGlyMetGlyLeuSerAlaLeuGlnGlnPheValGlyIle 282

Qy       984 CACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAA 1043
             :::   :::   :::||||||   ::::::::::::|||   |||   |||   :::|||
Db       283 AsnValIlePheTyrTyrSerSerValLeuTrpArgSerValGlyPheThrGluGluLys 302

Qy      1044 CACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCG 1103
               |||   |||            |||            |||      |||:::|||
Db       303 SerLeuLeuIleThrValIleThr---GlyPheIleAsnIleLeuThrThrLeuValAla 321

Qy      1104 ACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTG---GGCAGCACGGGC--- 1157
               |||:::|||      ||||||:::||||||||||||   ||||||   |||
Db       322 IleAlaPheValAspLysPheGlyArgLysProLeuLeuLeuMetGlySerIleGlyMet 341

Qy      1158 ------------GGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTC 1205
                         ||||||:::      :::   |||         |||   |||||||||
Db       342 ThrIleThrLeuGlyIleLeuSerValValPheGly---------GlyAlaThrValVal 358

Qy      1206 GGCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGC---CTAAGCATCGCCTCCACC 1262
                  |||          :::      |||   |||   :::::::::   :::
Db       359 AsnGlyGlnPro------ThrLeuThrGlyAlaAlaGlyIleIleAlaLeuValThrAla 376

Qy      1263 CTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCG 1322
               ||||||      |||             |||||||||   ||||||
Db       377 AsnLeuTyrValPheSerPheGlyPheSerTrpGlyProIleValTrpValLeuLeuGly 396

Qy      1323 GAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGC 1382
             |||:::|||      ::::::||||||          |||:::
Db       397 GluMetPheAsnAsnLysIleArgAlaAlaAlaLeuSerValAlaAlaGlyValGlnTrp 416

Qy      1383 GTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGC 1442
             :::   :::   :::||||||   ||||||      |||         :::   :::|||
Db       417 IleAlaAsnPheIleIleSerThrThrPheProProLeuLeuAspThrValGlyLeuGly 436

Qy      1443 GGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTAC 1502
```

```
                   ::::::   ||||||:::      ||||||::::::      |||   :::   :::
Db       437 ProAlaTyrGlyLeuTyrAlaThrSerAlaAlaIleSerIlePhePheIleTrpPhePhe 456

Qy      1503 CTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATG 1538
                :::  .||||||:::|||:::|||||||||:::|||
Db       457 ValLysGluThrLysGlyLysThrLeuGluGlnMet 468


Search completed: June 30, 2004, 18:44:40
Job time : 71.5 secs
```

GenCore version 5.1.6

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 30, 2004, 18:38:20 ; Search time 103.5 Seconds
(without alignments)
12297.589 Million cell updates/sec

Title: US-10-051-902A-21
Perfect score: 3614
Sequence: 1 cttacatgtaagctcgtgcc.....aaaaaaaaaaaaaaaaaaaa 2017

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 2034082

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10051902/runat_30062004_164723_20411/app_query.fasta_1.2183
-DB=SPTREMBL_25 -QFMT=fastan -SUFFIX=rspt -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10051902_@CGN_1_1_112_@runat_30062004_164723_20411 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2560	70.8	577	10	Q945E5	Q945e5 oryza sativ
2	1742	48.2	511	10	Q8LHC3	Q8lhc3 oryza sativ
3	1710	47.3	538	10	Q84QH3	Q84qh3 prunus cera
4	1625.5	45.0	539	10	Q9LS92	Q9ls92 arabidopsis
5	1617	44.7	523	10	Q7XA50	Q7xa50 glycine max
6	1607	44.5	509	10	Q84KI7	Q84ki7 prunus cera
7	1596.5	44.2	481	10	Q84RI1	Q84ri1 malus domes
8	1590	44.0	545	10	P93076	P93076 beta vulgar
9	1589	44.0	549	10	P93075	P93075 beta vulgar
10	1579	43.7	511	10	Q9XIH7	Q9xih7 arabidopsis
11	1559	43.1	511	10	Q9XIH6	Q9xih6 arabidopsis
12	1519	42.0	491	10	Q84RI2	Q84ri2 malus domes
13	1513	41.9	513	10	Q9FQX3	Q9fqx3 apium grave
14	1494	41.3	519	10	Q7XB36	Q7xb36 orobanche r
15	1398	38.7	523	10	Q8RVQ2	Q8rvq2 apium grave
16	1385.5	38.3	508	10	Q9ZNS0	Q9zns0 arabidopsis
17	1361	37.7	493	10	O23213	O23213 arabidopsis
18	1307.5	36.2	574	10	Q9AUM9	Q9aum9 oryza sativ
19	1304	36.1	506	10	Q8W2W8	Q8w2w8 oryza sativ
20	1304	36.1	506	10	Q7XFD6	Q7xfd6 oryza sativ
21	1186	32.8	538	10	Q7XKF1	Q7xkf1 oryza sativ
22	1167.5	32.3	535	10	Q7XKF0	Q7xkf0 oryza sativ
23	1143	31.6	547	10	Q9SKT9	Q9skt9 arabidopsis
24	1134.5	31.4	523	10	Q7X6M3	Q7x6m3 oryza sativ
25	673.5	18.6	521	10	O22848	O22848 arabidopsis
26	670.5	18.6	457	16	P96742	P96742 bacillus su
27	669.5	18.5	509	10	Q8VZR6	Q8vzr6 arabidopsis
28	653	18.1	467	16	Q8CQA7	Q8cqa7 staphylococ
29	649.5	18.0	596	10	Q7XIZ0	Q7xiz0 oryza sativ
30	643.5	17.8	498	10	Q84UY4	Q84uy4 mesembryant
31	639.5	17.7	618	11	Q921A2	Q921a2 rattus norv
32	630	17.4	245	10	Q8VZ80	Q8vz80 arabidopsis
33	629.5	17.4	582	10	O23492	O23492 arabidopsis
34	626	17.3	580	10	Q9C757	Q9c757 arabidopsis
35	623	17.2	493	10	Q7XRI6	Q7xri6 oryza sativ
36	618	17.1	581	10	Q9LKH1	Q9lkh1 mesembryant
37	612	16.9	409	16	Q83EH4	Q83eh4 coxiella bu
38	609.5	16.9	473	16	O34718	O34718 bacillus su
39	608.5	16.8	544	10	Q93WT7	Q93wt7 olea europa
40	608	16.8	558	10	Q9FIF2	Q9fif2 arabidopsis
41	600.5	16.6	525	10	Q852B0	Q852b0 oryza sativ
42	594	16.4	517	16	Q8G3X1	Q8g3x1 bifidobacte
43	591.5	16.4	546	10	Q93Z41	Q93z41 arabidopsis
44	591	16.4	581	10	Q7XQ00	Q7xq00 oryza sativ
45	589.5	16.3	479	2	Q9FDM0	Q9fdm0 zymomonas m

ALIGNMENTS

RESULT 1
Q945E5
ID Q945E5 PRELIMINARY; PRT; 577 AA.
AC Q945E5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative sugar transporter.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Wu P., Chen Q., Huang G., Yi K.;
RT "Molecular cloning of putative sugar transporter in rice.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL; AF416867; AAL14615.1; -.
DR Gramene; Q945E5; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.

```
DR   GO; GO:0008643; P:carbohydrate transport; IEA.
DR   InterPro; IPR007114; MFS.
DR   InterPro; IPR005828; Sub_transporter.
DR   InterPro; IPR003663; Sugar_transpt.
DR   InterPro; IPR005829; Sug_transporter.
DR   Pfam; PF00083; sugar_tr; 1.
DR   PRINTS; PR00171; SUGRTRNSPORT.
DR   TIGRFAMs; TIGR00879; SP; 1.
DR   PROSITE; PS50850; MFS; 1.
DR   PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR   PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW   Sugar transport; Transmembrane; Transport.
SQ   SEQUENCE   577 AA;  61178 MW;  95651621AE980CD4 CRC64;

Alignment Scores:
Pred. No.:              1.19e-141      Length:         577
Score:                  2560.00        Matches:        511
Percent Similarity:     97.36%         Conservative:   5
Best Local Similarity:  96.42%         Mismatches:     14
Query Match:            70.84%         Indels:         0
DB:                     10             Gaps:           0

US-10-051-902A-21 (1-2017) x Q945E5 (1-577)

Qy         36 TCGACCGCCACTACTGTACACGGCCCAGAGCGAGCCTCCTCCTCCTCTGCACCACCGGAG 95
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         48 SerThrAlaThrThrValHisGlyProGluArgAlaSerSerSerSerAlaProProGlu 67

Qy         96 ATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTC 155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 MetAlaSerAlaAlaLeuProGluAlaValAlaProLysLysLysGlyAsnValArgPhe 87

Qy        156 GCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGG 215
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         88 AlaPheAlaCysAlaIleLeuAlaSerMetThrSerIleLeuLeuGlyTyrAspIleGly 107

Qy        216 GTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTG 275
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 ValMetSerGlyAlaSerLeuTyrIleLysLysAspPheAsnIleSerAspGlyLysVal 127

Qy        276 GAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGG 335
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 GluValLeuMetGlyIleLeuAsnLeuTyrSerLeuIleGlySerPheAlaAlaGlyArg 147

Qy        336 ACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCG 395
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 ThrSerAspTrpIleGlyArgArgTyrThrIleValPheAlaAlaValIlePhePheAla 167

Qy        396 GGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTG 455
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 GlyAlaPheLeuMetGlyPheAlaValAsnTyrAlaMetLeuMetPheGlyArgPheVal 187

Qy        456 GCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCG 515
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 AlaGlyIleGlyValGlyTyrAlaLeuMetIleAlaProValTyrThrAlaGluValSer 207

Qy        516 CCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATC 575
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 ProAlaSerAlaArgGlyPheLeuThrSerPheProGluValPheIleAsnPheGlyIle 227

Qy        576 CTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGC 635
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        228 LeuLeuGlyTyrValSerAsnTyrAlaPheSerArgLeuProLeuAsnLeuGlyTrpArg 247

Qy        636 ATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATG 695
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 IleMetLeuGlyIleGlyAlaAlaProSerValLeuLeuAlaLeuMetValLeuGlyMet 267

Qy        696 CCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTG 755
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 ProGluSerProArgTrpLeuValMetLysGlyArgLeuAlaAspAlaLysValValLeu 287

Qy        756 GAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCC 815
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        288 GluLysThrSerAspThrAlaGluGluAlaAlaGluArgLeuAlaAspIleLysAlaAla 307

Qy        816 GCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGA 875
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 AlaGlyIleProGluGluLeuAspGlyAspValValThrValProLysArgGlySerGly 327

Qy        876 AACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATC 935
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 AsnGluLysArgValTrpLysGluLeuIleLeuSerProThrProAlaMetArgArgIle 347

Qy        936 CTGCTGTCCGGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTC 995
              ||||||||||||||||||||||||||||||||   |||   ||||||  |||||||||
Db        348 LeuLeuSerGlyIleGlyIleHisPhePheGlnGlnAlaSerGlyIleAspSerValVal 367

Qy        996 TTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGC 1055
                 |||||||||   ||||||||||||   |||:::|||:::|||||||||   ||||||
Db        368 LeuTyrSerProArgValPheLysSerAlaGlyIleThrAspAspLysHisLeuLeuGly 387

Qy       1056 ACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATC 1115
              ||||||          ||||||||||||   ||||||||||||:::||||||||||||:::
Db        388 ThrThrCysAlaValGlyValThrLysThrLeuPheIleLeuValAlaThrPhePheLeu 407

Qy       1116 GACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTC 1175
              |||   ||||||||||||||||||||||||||||   ||||||||||||:::||||||||||
Db        408 AspArgValGlyArgArgProLeuLeuLeuSerSerThrGlyGlyMetIleLeuSerLeu 427

Qy       1176 ATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTGG 1235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 IleGlyLeuGlyAlaGlyLeuThrValValGlyGlnHisProAspAlaLysIleProTrp 447

Qy       1236 GCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTT 1295
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 AlaIleGlyLeuSerIleAlaSerThrLeuAlaTyrValAlaPhePheSerIleGlyLeu 467

Qy       1296 GGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGC 1355
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        468 GlyProIleThrTrpValTyrSerSerGluIlePheProLeuGlnValArgAlaLeuGly 487

Qy       1356 TGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTG 1415
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        488 CysSerLeuGlyValAlaAlaAsnArgValThrSerGlyValIleSerMetThrPheLeu 507

Qy       1416 TCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCG 1475
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 SerLeuSerLysAlaIleThrIleGlyGlySerPhePheLeuTyrSerGlyIleAlaAla 527

Qy       1476 CTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAG 1535
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        528 LeuAlaTrpValPhePheTyrThrTyrLeuProGluThrArgGlyArgThrLeuGluGlu 547

Qy       1536 ATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAGCCAAGGAG 1595
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        548 MetSerLysLeuPheGlyAspThrAlaAlaAlaSerGluSerAspGluProAlaLysGlu 567

Qy       1596 AAGAAGAAGGTGGAAATGGCCGCCACTAAC 1625
              |||||||||||||||||||||||||||||
Db        568 LysLysLysValGluMetAlaAlaThrAsn 577

RESULT 2
Q8LHC3
ID   Q8LHC3      PRELIMINARY;      PRT;   511 AA.
AC   Q8LHC3;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Similar to myo-inositol transporter 2.
GN   P0458E05.27.
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Nipponbare;
RA   Sasaki T., Matsumoto T., Yamamoto K.;
RT   "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT   clone:P0458E05.";
RL   Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR   EMBL; AP004365; BAC05627.1; -.
DR   Gramene; Q8LHC3; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0005351; F:sugar porter activity; IEA.
DR   GO; GO:0005215; F:transporter activity; IEA.
DR   GO; GO:0008643; P:carbohydrate transport; IEA.
DR   InterPro; IPR007114; MFS.
DR   InterPro; IPR005828; Sub_transporter.
DR   InterPro; IPR003663; Sugar_transpt.
DR   InterPro; IPR005829; Sug_transporter.
DR   Pfam; PF00083; sugar_tr; 1.
DR   PRINTS; PR00171; SUGRTRNSPORT.
DR   TIGRFAMs; TIGR00879; SP; 1.
DR   PROSITE; PS50850; MFS; 1.
DR   PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR   PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW   Sugar transport; Transmembrane; Transport.
SQ   SEQUENCE   511 AA;  54606 MW;  ECF9573E23E3C625 CRC64;

Alignment Scores:
Pred. No.:                7.88e-94      Length:          511
Score:                    1742.00       Matches:         350
Percent Similarity:       82.68%        Conservative:    51
Best Local Similarity:    72.16%        Mismatches:      72
Query Match:              48.20%        Indels:          12
DB:                       10            Gaps:            4

US-10-051-902A-21 (1-2017) x Q8LHC3 (1-511)

Qy      210 ATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGG 269
Db       19 IleAsnTyrIleCysGlyAlaSerLeuPheIleLysGluAspMetLysIleThrAspVal 38

Qy      270 AAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCG 329
Db       39 GluValGluIleLeuLeuGlyIleLeuAsnLeuTyrSerLeuValGlySerPheAlaAla 58

Qy      330 GGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTC 389
Db       59 GlyArgThrSerAspTrpIleGlyArgArgLeuThrIleIleLeuAlaAlaValIlePhe 78

Qy      390 TTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGC 449
Db       79 PheValGlyAlaIleMetMetGlyLeuSerValAsnTyrProMetLeuMetAlaGlyArg 98

Qy      450 TTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAG 509
Db       99 PheValAlaGlyIleGlyValGlyTyrAlaPheMetIleAlaProValTyrThrAlaGlu 118

Qy      510 GTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTC 569
Db      119 ValSerProAlaSerSerArgGlyPheLeuThrSerPheProGluValPheIleAsnPhe 138

Qy      570 GGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGG 629
Db      139 GlyIleLeuLeuGlyTyrValSerAsnTyrAlaPheSerArgLeuArgLeuGlnLeuGly 158

Qy      630 TGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTC 689
Db      159 TrpArgLeuMetLeuGlyValGlyAlaAlaProSerValAlaLeuAlaLeuMetValLeu 178

Qy      690 GGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTG 749
Db      179 AlaMetProGluSerProArgTrpLeuValMetLysGlyArgLeuAlaAspAlaLysVal 198

Qy      750 GTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAG 809
Db      199 ValLeuGlyGluThrSerAspThrAlaGluGluAlaAlaThrArgLeuAlaGluIleLys 218

Qy      810 GCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGG 869
Db      219 GluAlaValAlaIleProAlaAspLeuAspGlyAspValValAlaValProLysArg--- 237

Qy      870 AGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGG 929
Db      238 AlaGlyGlyGluArgArgValTrpLysGluLeuIleLeuSerProThrProAlaValArg 257

Qy      930 CGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCC 989
Db      258 ArgValLeuLeuSerAlaLeuGlyIleHisPhePheGlnGlnSerSerGlyIleAspAla 277

Qy      990 GTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTC 1049
Db      278 ValValLeuTyrSerProArgValPheGlnSerAlaGlyIleThrAspLysAsnLysLeu 297

Qy     1050 TTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTC 1109
Db      298 LeuGlyThrThrCysAlaValGlyValThrLysThrLeuPheIleLeuValAlaThrPhe 317

Qy     1110 TTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTC 1169
Db      318 ThrLeuAspArgPheGlyArgArgProLeuLeuLeuAlaSerAlaGlyGlyMetIleAla 337

Qy     1170 TCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATA 1229
Db      338 ThrLeuValThrLeuGlyLeuGlyLeuThrValIleGlyGluAspAlaThrGlyGly--- 356

Qy     1230 CCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATC 1289
Db      357 GlyTrpAlaIleAlaValSerIleAlaSerIleLeuAlaPheValAlaPhePheSerIle 376

Qy     1290 GGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCG 1349
Db      377 GlyLeuGlyProIleThrTrpValTyrSerSerGluIlePheProLeuHisLeuArgAla 396

Qy     1350 CTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACC 1409
Db      397 LeuGlyCysAlaLeuGlyValGlyLeuAsnArgValThrSerGlyValIleSerMetThr 416

Qy     1410 TTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATC 1469
Db      417 PheLeuSerLeuSerLysAlaIleThrIleGlyGlySerPhePheLeuTyrAlaGlyVal 436

Qy     1470 GCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTG 1529
Db      437 AlaSerLeuAlaTrpLeuPhePhePheThrTyrLeuProGluThrArgGlyArgThrLeu 456

Qy     1530 GAGGAGATGAGCAAGCTGTTC----------------------GGCGACACGGCCGCCGCC 1568
Db      457 GluGlnMetGlyGluLeuPheArgIleHisAsnMetAlaGlyAspAspAspSerAlaAla 476

Qy     1569 TCGGAATCAGACGAGCCAGCCAAGGAGAAGAAG---------AAGGTGGAAATGGCCGCC 1619
Db      477 ThrArgProProSerProGluGluGluGluLysIleThrAsnTyrValGluMetAlaAla 496

Qy     1620 ACTAACTGATCAAAC 1634
Db      497 ProSerSerSerSer 501

RESULT 3
Q84QH3
ID   Q84QH3      PRELIMINARY;      PRT;   538 AA.
AC   Q84QH3;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Putative sorbitol transporter.
```

```
GN   SORT2.
OS   Prunus cerasus.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.
OX   NCBI_TaxID=140311;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Montmorency; TISSUE=Fruit;
RX   MEDLINE=22578918; PubMed=12692316;
RA   Gao Z., Maurousset L., Lemoine R., Yoo S.D., Van Nocker S.,
RA   Loescher W.;
RT   "Cloning, Expression, and Characterization of Sorbitol Transporters
RT   from Developing Sour Cherry Fruit and Leaf Sink Tissues.";
RL   Plant Physiol. 131:1566-1575(2003).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Montmorency; TISSUE=Fruit;
RA   Zhifang G., Loescher W.H.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY100638; AAM44082.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005351; F:sugar porter activity; IEA.
DR   GO; GO:0005215; F:transporter activity; IEA.
DR   GO; GO:0008643; P:carbohydrate transport; IEA.
DR   InterPro; IPR007114; MFS.
DR   InterPro; IPR005828; Sub_transporter.
DR   InterPro; IPR003663; Sugar_transpt.
DR   InterPro; IPR005829; Sug_transporter.
DR   Pfam; PF00083; sugar_tr; 1.
DR   PRINTS; PR00171; SUGRTRNSPORT.
DR   TIGRFAMs; TIGR00879; SP; 1.
DR   PROSITE; PS50850; MFS; 1.
DR   PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR   PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
SQ   SEQUENCE   538 AA;  58569 MW;  ADDD03B983C6D358 CRC64;
```

Alignment Scores:

Pred. No.:	5.91e-92	Length:	538
Score:	1710.00	Matches:	341
Percent Similarity:	76.76%	Conservative:	62
Best Local Similarity:	64.95%	Mismatches:	114
Query Match:	47.32%	Indels:	8
DB:	10	Gaps:	4

US-10-051-902A-21 (1-2017) x Q84QH3 (1-538)

```
Qy         81 TCTGCACCACCGGAGATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAG 140
Db         13 SerGlyGlnProGlnLysAsnIleAlaAspPheAspProProGlyLysProLysArgAsn 32

Qy        141 GGCAACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTC 200
Db         33 ---------LysTyrAlaPheAlaCysAlaIleLeuAlaSerMetThrSerIleLeuLeu 49

Qy        201 GGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATC 260
Db         50 GlyTyrAspIleGlyValMetSerGlyAlaValIleTyrIleLysLysAspLeuLysVal 69

Qy        261 AGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCC 320
Db         70 SerAspValGluIleGluValLeuValGlyIleLeuAsnLeuTyrSerLeuIleGlySer 89

Qy        321 TTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCC 380
Db         90 AlaAlaAlaGlyArgThrSerAspTrpIleGlyArgArgTyrThrIleValLeuAlaGly 109

Qy        381 GTCATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATG 440
Db        110 AlaIlePhePheAlaGlyAlaLeuLeuMetGlyPheAlaProAsnTyrAlaPheLeuMet 129

Qy        441 TTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTAC 500
```

```
Db        130 PheGlyArgPheValAlaGlyIleGlyValGlyTyrAlaLeuMetIleAlaProValTyr 149

Qy        501 ACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTC 560
Db        150 ThrAlaGluValSerProAlaSerSerArgGlyPheLeuThrSerProGluValPhe 169

Qy        561 ATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTG 620
Db        170 IleAsnAlaGlyIleLeuPheGlyTyrValSerAsnTyrGlyPheSerLysLeuProThr 189

Qy        621 AACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTC 680
Db        190 HisLeuGlyTrpArgLeuMetLeuGlyValGlyAlaIleProSerIlePheLeuAlaIle 209

Qy        681 ATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGAC 740
Db        210 GlyValLeuAlaMetProGluSerProArgTrpLeuValMetGlnGlyArgLeuGlyAsp 229

Qy        741 GCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCC 800
Db        230 AlaArgLysValLeuAspLysThrSerAspSerLeuGluGluSerLysLeuArgLeuGly 249

Qy        801 GACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCC 860
Db        250 GluIleLysGluAlaAlaGlyIleProGluHisCysAsnAspAspIleValGluValLys 269

Qy        861 AAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCG 920
Db        270 LysArgSerGlnGlyGlnGlu------ValTrpLysGlnLeuLeuLeuArgProThrPro 287

Qy        921 GCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGC 980
Db        288 AlaValArgHisIleLeuMetCysAlaValGlyLeuHisPhePheGlnGlnAlaSerGly 307

Qy        981 ATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGAC 1040
Db        308 IleAspAlaValValLeuTyrSerProArgIlePheGluLysAlaGlyIleThrAsnPro 327

Qy       1041 AAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTG 1100
Db        328 AspHisValLeuLeuCysThrValAlaValGlyPheValLysThrValPheIleLeuVal 347

Qy       1101 GCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGG 1160
Db        348 AlaThrPheMetLeuAspArgIleGlyArgArgProLeuLeuLeuThrSerValAlaGly 367

Qy       1161 ATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGAC 1220
Db        368 MetValPheThrLeuAlaCysLeuGlyLeuGlyLeuThrIleIle---AspHisSerGly 386

Qy       1221 GCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTC 1280
Db        387 GluLysIleMetTrpAlaIleAlaLeuSerLeuThrMetValLeuAlaTyrValAlaPhe 406

Qy       1281 TTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAG 1340
Db        407 PheSerIleGlyMetGlyProIleThrTrpValTyrSerSerGluIlePheProLeuGln 426

Qy       1341 GTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATC 1400
Db        427 LeuArgAlaGlnGlyCysSerIleGlyValAlaValAsnArgValValSerGlyValLeu 446

Qy       1401 TCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTAC 1460
Db        447 SerMetThrPheIleSerLeuTyrLysAlaIleThrIleGlyGlyAlaPhePheLeuPhe 466

Qy       1461 TCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGC 1520
Db        467 AlaAlaIleAlaAlaValGlyTrpThrPhePhePheThrMetLeuProGluThrGlnGly 486

Qy       1521 CGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGAC 1580
Db        487 ArgThrLeuGluAspMetGluValLeuPheGlyLysPheTyrArgTrpArgLysAlaAsn 506
```

```
Qy       1581 GAGCCAGCCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCACTAACTGATCAAACTAACCG 1640
                 |||:::||||||:::|||:::             |||      |||   |||
Db        507 AlaLeuLeuLysGlnLysLysGlnValAspHisGlyAspGlyAsn------AsnAsnPro 524

Qy       1641 CAAAATCACCAAATC 1655
                 |||   ||||||
Db        525 AsnAsnProGlnIle 529
```

RESULT 4
Q9LS92
ID Q9LS92 PRELIMINARY; PRT; 539 AA.
AC Q9LS92;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sugar transporter protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
RL Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RX MEDLINE=20277480; PubMed=10819329;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT features of the regions of 4,504,864 bp covered by sixty P1 and TAC
RT clones.";
RL DNA Res. 7:131-135(2000).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL; AB026654; BAB01812.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0008643; P:carbohydrate transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00083; sugar_tr; 1.
DR PRINTS; PR00171; SUGRTRNSPORT.
DR TIGRFAMs; TIGR00879; SP; 1.
DR PROSITE; PS50850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Sugar transport; Transmembrane; Transport.
SQ SEQUENCE 539 AA; 58102 MW; 36B389284E5563A0 CRC64;

Alignment Scores:

Pred. No.:	5.16e-87	Length:	539
Score:	1625.50	Matches:	328
Percent Similarity:	75.76%	Conservative:	69
Best Local Similarity:	62.60%	Mismatches:	116
Query Match:	44.98%	Indels:	11
DB:	10	Gaps:	6

US-10-051-902A-21 (1-2017) x Q9LS92 (1-539)

```
Qy         60 CCAGAGCGAGCCTCCTCCTCCTCTGCACCACCGGAGATGGCTTCCGCCGCGCTGCCGGAG 119
              ||||||      ::::::   |||  ||||||                  :::||||||
Db          6 ProGluAsnArgThrAlaProSerProProPro---------ValLysHisValProGlu 22

Qy        120 GCCGTCGCGCCGAAGAAG------AAGGGCAACGTCCGGTTCGCCTTCGCCTGCGCCATC 173
              :::|||   |||   |||      |||   |||      :::||||||||||||||||||
Db         23 SerValLeuProAlaLysProProLysArgAsn---AsnTyrAlaPheAlaCysAlaIle 41

Qy        174 CTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCG 233
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         42 LeuAlaSerMetThrSerIleLeuLeuGlyTyrAspIleGlyValMetSerGlyAlaMet 61

Qy        234 CTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATA 293
              :::|||||||||:::|||      |||:::|||   ::::::   :::|||   |||
Db         62 IleTyrIleLysArgAspLeuLysIleAsnAspLeuGlnIleGlyIleLeuAlaGlySer 81

Qy        294 CTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGGCGGACGTCGGACTGGATCGGC 353
              ||||||:::|||||||||||||||||||   ||||||||||||||||||||||||||||
Db         82 LeuAsnIleTyrSerLeuIleGlySerCysAlaAlaGlyArgThrSerAspTrpIleGly 101

Qy        354 CGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGSGTTCCTCATGGGG 413
              |||||||||||||||||   |||      ||||||||||||||      |||||||||
Db        102 ArgArgTyrThrIleValLeuAlaGlyAlaIlePhePheAlaGlyAlaIleLeuMetGly 121

Qy        414 TTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGC 473
                 :::   ||||||||   ||||||||||||||||:::|||||||||||||||||||
Db        122 LeuSerProAsnTyrAlaPheLeuMetPheGlyArgPheIleAlaGlyIleGlyValGly 141

Qy        474 TACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGC 533
              ||||||||||||||||||||||||||||||||||||||||||||||||||:::||||||
Db        142 TyrAlaLeuMetIleAlaProValTyrThrAlaGluValSerProAlaSerSerArgGly 161

Qy        534 TTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCG 593
              ||||||   |||||||||||||||||||||||   ||||||:::|||||||||||||||
Db        162 PheLeuAsnSerPheProGluValPheIleAsnAlaGlyIleMetLeuGlyTyrValSer 181

Qy        594 AACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGC 653
              |||   ||||||||   ||||||||   :::||||||||:::|||||||||||||||||
Db        182 AsnLeuAlaPheSerAsnLeuProLeuLysValGlyTrpArgLeuMetLeuGlyIleGly 201

Qy        654 GCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGG 713
              |||   ||||||||:::||||||:::   ||||||   ||||||||||||||||||||
Db        202 AlaValProSerValIleLeuAlaIleGlyValLeuAlaMetProGluSerProArgTrp 221

Qy        714 CTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACG 773
              |||||||||:::|||||||||   ||||||||   ||||||:::|||||||||||||:::
Db        222 LeuValMetGlnGlyArgLeuGlyAspAlaLysArgValLeuAspLysThrSerAspSer 241

Qy        774 GCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAG 833
                  ||||||      ||||||   |||||||||   ||||||||||||||||   :::
Db        242 ProThrGluAlaThrLeuArgLeuGluAspIleLysHisAlaAlaGlyIleProAlaAsp 261

Qy        834 CTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGG 893
                    |||||||||   |||   :::|||   |||      |||      ||||||
Db        262 CysHisAspAspValValGlnValSerArgArgAsnSerHisGlyGlu---GlyValTrp 280

Qy        894 AAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGC 953
              :::||||||::::::   |||||||||||:::||||||:::::::::::::   ||||||
Db        281 ArgGluLeuLeuIleArgProThrProAlaValArgArgValMetIleAlaAlaIleGly 300

Qy        954 ATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTG 1013
              ||||||||||||||   |||   ||||||   :::||||||   :::||||||   :::
Db        301 IleHisPhePheGlnGlnAlaSerGlyIleAspAlaValValLeuPheSerProArgIle 320

Qy       1014 TTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGT 1073
              ||||||:::   ||||||      |||         |||      |||         |||
Db        321 PheLysThrAlaGlyLeuLysThrAspHisGlnGlnLeuLeuAlaThrValAlaValGly 340

Qy       1074 GTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGG 1133
              |||   |||      |||||||||:::||||||||   :::|||   :::|||||||||
Db        341 ValValLysThrSerPheIleLeuValAlaThrPheLeuLeuAspArgIleGlyArgArg 360

Qy       1134 CCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGG 1193
              ||||||||||||   |||   ||||||::::::|||||||||      ||||||
Db        361 ProLeuLeuLeuThrSerValGlyGlyMetValLeuSerLeuAlaAlaLeuGlyThrSer 380

Qy       1194 CTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATC 1253
              ||||||::::::   |||      :::   |||:::   ||||||:::   ::::::|||
Db        381 LeuThrIleIleAspGln---SerGluLysLysValMetTrpAlaValValValAlaIle 399
```

```
Qy       1254 GCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTG 1313
Db        400 AlaThrValMetThrTyrValAlaThrPheSerIleGlyAlaGlyProIleThrTrpVal 419

Qy       1314 TACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCC 1373
Db        420 TyrSerSerGluIlePheProLeuArgLeuArgSerGlnGlySerSerMetGlyValVal 439

Qy       1374 GCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATC 1433
Db        440 ValAsnArgValThrSerGlyValIleSerIleSerPheLeuProMetSerLysAlaMet 459

Qy       1434 ACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTC 1493
Db        460 ThrThrGlyGlyAlaPheTyrLeuPheGlyGlyIleAlaThrValAlaTrpValPhePhe 479

Qy       1494 TACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGC 1553
Db        480 TyrThrPheLeuProGluThrGlnGlyArgMetLeuGluAspMetAspGluLeuPhe--- 498

Qy       1554 GACACGGCCGCCGCCTCGGAATCAGACGAGCCAGCCAAGGAGAAGAAGAAGGTGGAAATG 1613
Db        499 ------SerGlyPheArgTrpArgAspSerLysSerLysProLysGlyAsnProGluLys 516

Qy       1614 GCCGCCACTAAC 1625
Db        517 ThrValProAsn 520
```

RESULT 5
Q7XA50
ID Q7XA50 PRELIMINARY; PRT; 523 AA.
AC Q7XA50;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sorbitol-like transporter.
GN STP.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Williams;
RA Dimou M., Aivalakis G., Katinakis P.;
RT "Carbon metabolism in developing roots and lateral roots of etiolated
RT Glycine max seedlings. Expression of GmSTP gene coding for a sorbitol-
RT like transporter.";
RL Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AJ563367; CAD91337.1; -.
SQ SEQUENCE 523 AA; 56489 MW; 3A586397262F1C4B CRC64;

Alignment Scores:

Pred. No.:	1.61e-86	Length:	523
Score:	1617.00	Matches:	325
Percent Similarity:	74.43%	Conservative:	68
Best Local Similarity:	61.55%	Mismatches:	117
Query Match:	44.74%	Indels:	18
DB:	10	Gaps:	4

US-10-051-902A-21 (1-2017) x Q7XA50 (1-523)

```
Qy         99 GCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCC 158
Db         12 AlaHisLysThrLeuGlnAspPheAspProProLysLysArgLysArgAsnLysTyrAla 31

Qy        159 TTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTG 218
Db         32 PheAlaCysAlaMetLeuAlaSerMetThrSerIleLeuLeuGlyTyrAspIleGlyVal 51

Qy        219 ATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAG 278
```

```
Db         52 MetSerGlyAlaAlaIleTyrIleLysArgAspLeuLysValSerAspGluGlnIleGlu 71

Qy        279 GTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACG 338
Db         72 IleLeuLeuGlyIleIleAsnLeuTyrSerLeuIleGlySerCysLeuAlaGlyArgThr 91

Qy        339 TCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGG 398
Db         92 SerAspTrpIleGlyArgArgTyrThrIleGlyLeuGlyGlyAlaIlePheLeuValGly 111

Qy        399 GSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCC 458
Db        112 SerThrLeuMetGlyPheTyrProHisTyrSerPheLeuMetCysGlyArgPheValAla 131

Qy        459 GGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCG 518
Db        132 GlyIleGlyIleGlyTyrAlaLeuMetIleAlaProValTyrThrAlaGluValSerPro 151

Qy        519 GCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTG 578
Db        152 AlaSerSerArgGlyPheLeuThrSerProGluValPheIleAsnGlyGlyIleLeu 171

Qy        579 CTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATC 638
Db        172 LeuGlyTyrIleSerAsnTyrGlyPheSerLysLeuThrLeuLysValGlyTrpArgMet 191

Qy        639 ATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCG 698
Db        192 MetLeuGlyValGlyAlaIleProSerValValLeuThrGluGlyValLeuAlaMetPro 211

Qy        699 GAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAG 758
Db        212 GluSerProArgTrpLeuValMetArgGlyArgLeuGlyGluAlaArgLysValLeuAsn 231

Qy        759 AAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCC 818
Db        232 LysThrSerAspSerLysGluGluAlaGlnLeuArgLeuAlaGluIleLysGlnAlaAla 251

Qy        819 GGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAAC 878
Db        252 GlyIleProGluSerCysAsnAspAspValValGlnValAsnLysGlnSerAsnGly--- 270

Qy        879 GAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTG 938
Db        271 ---GluGlyValTrpLysGluLeuPheLeuTyrProThrProAlaIleArgHisIleVal 289

Qy        939 CTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTC 998
Db        290 IleAlaAlaLeuGlyIleHisPhePheGlnGlnAlaSerGlyValAspAlaValValLeu 309

Qy        999 TACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGCACC 1058
Db        310 TyrSerProArgIlePheGluLysAlaGlyIleThrAsnAspThrHisLysLeuLeuAla 329

Qy       1059 ACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGAC 1118
Db        330 ThrValAlaValGlyPheValLysThrValPheIleLeuAlaAlaThrPheThrLeuAsp 349

Qy       1119 GGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATC 1178
Db        350 ArgValGlyArgArgProLeuLeuLeuSerSerValGlyGlyMetValLeuSerLeuLeu 369

Qy       1179 GGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCC 1238
Db        370 ThrLeuAlaIleSerLeuThrValIle---AspHisSerGluArgLysLeuMetTrpAla 388

Qy       1239 ATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGC 1298
Db        389 ValGlySerSerIleAlaMetValLeuAlaTyrValAlaThrPheSerIleGlyAlaGly 408

Qy       1299 CCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGC 1358
Db        409 ProIleThrTrpValTyrSerSerGluIlePheProLeuArgLeuArgAlaGlnGlyAla 428
```

```
Qy       1359 TCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCG 1418
Db        429 AlaAlaGlyValAlaValAsnArgThrThrSerAlaValValSerMetThrPheLeuSer 448

Qy       1419 CTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTC 1478
Db        449 LeuThrArgAlaIleThrIleGlyGlyAlaPhePheLeuTyrCysGlyIleAlaThrVal 468

Qy       1479 GCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATG 1538
Db        469 GlyTrpIlePhePheTyrThrValLeuProGluThrArgGlyLysThrLeuGluAspMet 488

Qy       1539 AGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAGCCAAGGAGAAG 1598
Db        489 GluGlySerPheGlyThrPheArgSerLysSerAsnAla----------------Ser 502

Qy       1599 AAGAAGGTGGAAATGGCCGCCACTAACTGATCAAACTAACCGCAAAATCACCAAATCCTA 1658
Db        503 LysAlaValGlu------------------------AsnGluAsnGlyGlnValAla 513

Qy       1659 AGGGTTTTCTTGCAAAAACGTGTG 1682
Db        514 GlnValGlnLeuGlyThrAsnVal 521

RESULT 6
Q84KI7
ID   Q84KI7      PRELIMINARY;      PRT;   509 AA.
AC   Q84KI7;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Sorbitol transporter.
OS   Prunus cerasus.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.
OX   NCBI_TaxID=140311;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fruit;
RX   MEDLINE=22578918; PubMed=12692316;
RA   Gao Z., Maurousset L., Lemoine R., Yoo S.-D., van Nocker S.,
RA   Loescher W.;
RT   "Cloning, Expression, and Characterization of Sorbitol Transporters
RT   from Developing Sour Cherry Fruit and Leaf Sink Tissues.";
RL   Plant Physiol. 131:1566-1575(2003).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fruit;
RA   Zhifang G., Loescher W.H., Lemoine R.;
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF482011; AAO39267.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005351; F:sugar porter activity; IEA.
DR   GO; GO:0005215; F:transporter activity; IEA.
DR   GO; GO:0008643; P:carbohydrate transport; IEA.
DR   InterPro; IPR007114; MFS.
DR   InterPro; IPR005828; Sub_transporter.
DR   InterPro; IPR003663; Sugar_transpt.
DR   InterPro; IPR005829; Sug_transporter.
DR   Pfam; PF00083; sugar_tr; 1.
DR   PRINTS; PR00171; SUGRTRNSPORT.
DR   TIGRFAMs; TIGR00879; SP; 1.
DR   PROSITE; PS50850; MFS; 1.
DR   PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.
DR   PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
SQ   SEQUENCE   509 AA;  55407 MW;  EA25ED065CE884C8 CRC64;

Alignment Scores:
Pred. No.:              6.17e-86      Length:          509
Score:                  1607.00       Matches:         319
Percent Similarity:     78.48%        Conservative:    53
Best Local Similarity:  67.30%        Mismatches:      98
Query Match:            44.47%        Indels:          4
DB:                     10            Gaps:            3

US-10-051-902A-21 (1-2017) x Q84KI7 (1-509)

Qy        132 AAGAAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCC 191
Db         20 LysProLysArgAsnLeu---TyrAlaIleGlyCysAlaIleLeuAlaSerMetThrSer 38

Qy        192 ATCCTCCTCGGCTACGATATCGGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGAC 251
Db         39 IleLeuLeuGlyTyrAspIleGlyValMetSerGlyAlaSerIleTyrIleGlnLysAsp 58

Qy        252 TTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTC 311
Db         59 LeuLysIleSerAspValGluValGluIleLeuIleGlyIleLeuAsnLeuTyrSerLeu 78

Qy        312 ATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTG 371
Db         79 IleGlySerAlaAlaAlaGlyArgThrSerAspTrpIleGlyArgArgTyrThrIleVal 98

Qy        372 TTCGCCGCCGTCATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCC 431
Db         99 PheAlaGlyAlaIlePhePheThrGlyAlaLeuLeuMetGlyLeuAlaThrAsnTyrAla 118

Qy        432 ATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCG 491
Db        119 PheLeuMetValGlyArgPheValAlaGlyIleGlyValGlyTyrAlaLeuMetIleAla 138

Qy        492 CCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCG 551
Db        139 ProValTyrAsnAlaGluValSerProAlaSerSerArgGlyAlaLeuThrSerPhePro 158

Qy        552 GAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGC 611
Db        159 GluValPheValAsnIleGlyIleLeuLeuGlyTyrValAlaAsnTyrAlaPheSerGly 178

Qy        612 TTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTG 671
Db        179 LeuProIleAspLeuGlyTrpArgLeuMetLeuGlyValGlyValPheProSerValIle 198

Qy        672 CTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGC 731
Db        199 LeuAlaValGlyValLeuSerMetProGluSerProArgTrpLeuValMetGlnGlyArg 218

Qy        732 CTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAG 791
Db        219 LeuGlyGluAlaLysGlnValLeuAspLysThrSerAspSerLeuGluGluAlaGlnLeu 238

Qy        792 CGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTG 851
Db        239 ArgLeuAlaAspIleLysGluAlaAlaGlyIleProGluHisCysValGluAspValVal 258

Qy        852 ACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCC 911
Db        259 GlnValProLysHisSerHisGlyGluGlu------ValTrpLysGluLeuLeuLeuHis 276

Qy        912 CCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCAT 971
Db        277 ProThrProProValArgHisIleLeuIleAlaAlaIleGlyPheHisPhePheGlnGln 296

Qy        972 GCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTA 1031
Db        297 LeuSerGlyIleAspAlaLeuValLeuTyrSerProArgIlePheGluLysAlaGlyIle 316

Qy       1032 ACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTC 1091
Db        317 ThrAspSerSerThrLeuLeuLeuAlaThrValAlaValGlyPheSerLysThrIlePhe 336

Qy       1092 ATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGC 1151
Db        337 ThrLeuValAlaIleGlyPheLeuAspArgValGlyArgArgProLeuLeuLeuThrSer 356
```

```
Qy       1152 ACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAG 1211
Db        357 ValAlaGlyMetIleAlaSerLeuLeuCysLeuGlyThrSerLeuThrIleVal---Asp 375

Qy       1212 CACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTAC 1271
Db        376 HisGluThrGluLysMetMetTrpAlaSerValLeuCysLeuThrMetValLeuAlaTyr 395

Qy       1272 GTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTC 1331
Db        396 ValGlyPhePheSerIleGlyMetGlyProIleAlaTrpValTyrSerSerGluIlePhe 415

Qy       1332 CCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGC 1391
Db        416 ProLeuLysLeuArgAlaGlnGlyCysSerMetGlyThrAlaValAsnArgIleMetSer 435

Qy       1392 GGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTC 1451
Db        436 GlyValLeuSerMetSerPheIleSerLeuTyrLysAlaIleThrMetGlyGlyThrPhe 455

Qy       1452 TTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAG 1511
Db        456 PheLeuTyrAlaGlyIleAlaThrValGlyTrpValPhePheTyrThrMetLeuProGlu 475

Qy       1512 ACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGC 1553
Db        476 ThrGlnGlyArgThrLeuGluAspMetGluValLeuPheGly 489
```

RESULT 7
Q84RI1
ID Q84RI1 PRELIMINARY; PRT; 481 AA.
AC Q84RI1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sorbitol transporter.
GN SOT2.
OS Malus domestica (Apple) (Malus sylvestris).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Rosales; Rosaceae; Maloideae; Malus.
OX NCBI_TaxID=3750;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Mutsu; TISSUE=Fruit;
RA Zhifang G., Jayanty S., Beaudry R.M., Loescher W.H.;
RT "Cloning and characterization of apple fruit sorbitol transporter.";
RL Submitted (FEB-2003) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY237401; AAO88965.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0008643; P:carbohydrate transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00083; sugar_tr; 1.
DR PRINTS; PR00171; SUGRTRNSPORT.
DR TIGRFAMs; TIGR00879; SP; 1.
DR PROSITE; PS50850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
SQ SEQUENCE 481 AA; 52566 MW; A3C20EE544A85C6E CRC64;

Alignment Scores:

Pred. No.:	2.51e-85	Length:	481
Score:	1596.50	Matches:	322
Percent Similarity:	78.78%	Conservative:	53
Best Local Similarity:	67.65%	Mismatches:	98
Query Match:	44.18%	Indels:	3
DB:	10	Gaps:	2

US-10-051-902A-21 (1-2017) x Q84RI1 (1-481)

```
Qy        183 ATGACCTCCATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACATC 242
Db          1 MetThrSerIleLeuLeuGlyTyrAspIleGlyValMetSerGlyAlaSerLeuPheIle 20

Qy        243 AAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTC 302
Db         21 LysGluAsnLeuLysIleSerAspValGlnValGluIleMetAsnGlyThrLeuAsnLeu 40

Qy        303 TACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTAC 362
Db         41 TyrSerLeuIleGlySerAlaLeuAlaGlyArgThrSerAspTrpIleGlyArgArgTyr 60

Qy        363 ACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGGSGTTCCTCATGGGGTTCGCCGTC 422
Db         61 ThrIleValLeuAlaGlyThrIlePhePheIleGlyAlaLeuLeuMetGlyPheAlaPro 80

Qy        423 AACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTC 482
Db         81 AsnTyrAlaPheLeuMetPheGlyArgPheValAlaGlyValGlyValGlyTyrAlaLeu 100

Qy        483 ATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACG 542
Db        101 MetIleAlaProValTyrThrAlaGluIleSerProAlaSerPheArgGlyPheLeuThr 120

Qy        543 TCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCT 602
Db        121 SerPheProGluValPheValAsnIleGlyIleLeuLeuGlyTyrValSerAsnTyrAla 140

Qy        603 TTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCG 662
Db        141 PheSerLysLeuProIleHisLeuAsnTrpArgIleMetLeuGlyValGlyAlaPhePro 160

Qy        663 TCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATG 722
Db        161 SerValIleLeuAlaValGlyValLeuAlaMetProGluSerProArgTrpLeuValMet 180

Qy        723 AAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAG 782
Db        181 GlnGlyArgLeuGlyAspAlaLysArgValLeuGlnLysThrSerGluSerIleGluGlu 200

Qy        783 GCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGC 842
Db        201 CysGlnLeuArgLeuAspAspIleLysGluAlaAlaGlyIleProLysGluSerAsnAsp 220

Qy        843 GACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTC 902
Db        221 AspValValGlnValSerLysArgSerHisGly------GluGlyValTrpLysGluLeu 238

Qy        903 ATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTC 962
Db        239 LeuLeuHisProThrProAlaValArgHisIleLeuIleAlaAlaLeuGlyIleHisPhe 258

Qy        963 TTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGC 1022
Db        259 PheGluGlnSerSerGlyIleAspSerValValLeuTyrSerProArgIlePheGluLys 278

Qy       1023 CCCGGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAG 1082
Db        279 AlaGlyIleThrSerTyrAspHisLysLeuLeuAlaThrValAlaValGlyValValLys 298

Qy       1083 AGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTG 1142
Db        299 ThrIleCysIleLeuValAlaThrValPheLeuAspLysPheGlyArgArgProLeuLeu 318

Qy       1143 CTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTC 1202
Db        319 LeuThrSerValAlaGlyMetValPheSerLeuSerCysLeuGlyAlaSerLeuThrIle 338

Qy       1203 GTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACC 1262
```

```
Db        339 ValAspGlnGln---HisGlyLysIleMetTrpAlaIleValLeuCysIleThrMetVal 357

Qy       1263 CTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCG 1322
              |||     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 LeuLeuAsnValAlaPhePheSerIleGlyLeuGlyProIleThrTrpValTyrSerSer 377

Qy       1323 GAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGC 1382
              |||||||||||||||||:::||||||   |||||||||:::|||||||||   ||||||
Db        378 GluIlePheProLeuGlnLeuArgAlaGlnGlyCysSerMetGlyValAlaValAsnArg 397

Qy       1383 GTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGC 1442
              ||||||||||||||||||||||||||||:::||||||   ||||||||||||||||||
Db        398 ValThrSerGlyValIleSerMetThrPheIleSerLeuTyrLysAlaIleThrIleGly 417

Qy       1443 GGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTAC 1502
              |||:::|||||||||||||:::||||||||||||:::   |||||||||||||||
Db        418 GlyAlaPhePheLeuTyrAlaGlyIleAlaAlaValGlyTrpValPhePheTyrMetLeu 437

Qy       1503 CTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGGCC 1562
                 |||||||||:::|||||||||||||||:::|||      |||||||||
Db        438 TyrProGluThrGlnGlyArgThrLeuGluAspMetGluValLeuPheGlyLysTyrHis 457

Qy       1563 GCCGCCTCGGAATCAGACGAGCCAGCCAAGGAGAAGAAGAAGGTGGAA 1610
                    |||::::::      ::::::   |||:::|||:::
Db        458 LysTrpArgGluAlaAsnAlaLeuGlnLysAlaLysGlnValAsp 473

RESULT 8
P93076
ID   P93076      PRELIMINARY;      PRT;   545 AA.
AC   P93076;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   BvcDNA-397 protein.
GN   BVCDNA-397.
OS   Beta vulgaris (Sugar beet).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Caryophyllales; Amaranthaceae; Beta.
OX   NCBI_TaxID=161934;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97198558; PubMed=9046601;
RA   Chiou T.-J., Bush D.R.;
RT   "Isolation and molecular characteristics of two putative sugar
RT   transporters from sugar beet (Accession Nos. U64902 and U64903)
RT   (PGR97-017).";
RL   Plant Physiol. 113:663-663(1997).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR   EMBL; U64903; AAB68029.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0005351; F:sugar porter activity; IEA.
DR   GO; GO:0005215; F:transporter activity; IEA.
DR   GO; GO:0008643; P:carbohydrate transport; IEA.
DR   InterPro; IPR007114; MFS.
DR   InterPro; IPR005828; Sub_transporter.
DR   InterPro; IPR003663; Sugar_transpt.
DR   InterPro; IPR005829; Sug_transporter.
DR   Pfam; PF00083; sugar_tr; 1.
DR   PRINTS; PR00171; SUGRTRNSPORT.
DR   TIGRFAMs; TIGR00879; SP; 1.
DR   PROSITE; PS50850; MFS; 1.
DR   PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR   PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW   Sugar transport; Transmembrane; Transport.
SQ   SEQUENCE   545 AA;  59321 MW;  9AD0290DB456BDAE CRC64;

Alignment Scores:
Pred. No.:              6.15e-85      Length:          545
Score:                  1590.00       Matches:         327
Percent Similarity:     72.07%        Conservative:    60
Best Local Similarity:  60.89%        Mismatches:      128

Query Match:            44.00%        Indels:          22
DB:                     10            Gaps:            5

US-10-051-902A-21 (1-2017) x P93076 (1-545)

Qy         78 TCCTCTGCACCACCGGAGATGGCTTCCGCCGCGCTG-----------CCGGAGGCCGTC 125
              |||     ||||||     ||||||     :::          |||
Db         10 SerAspProProProThrThrAlaSerLysValIleAlaAspPheAspProLeuLysLys 29

Qy        126 GCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATG 185
                 ||||||:::          :::|||||||||||||||||    ||||||||||||
Db         30 ProProLysArgAsn---------LysPheAlaPheAlaCysAlaThrLeuAlaSerMet 46

Qy        186 ACCTCCATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAG 245
              ||||||:::|||||||||||||||||||||||||||||||||   :::|||:::|||
Db         47 ThrSerValLeuLeuGlyTyrAspIleGlyValMetSerGlyAlaIleIleTyrLeuLys 66

Qy        246 AAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTAC 305
              :::|||:::::|||||||||   ::::::   ||||||:::|||||||||||:::|||
Db         67 GluAspTrpHisIleSerAspThrGlnIleGlyValLeuValGlyIleLeuAsnIleTyr 86

Qy        306 TCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACC 365
                 |||  |||||||||||||||||||||||||||||||||||||||||||||||||
Db         87 CysLeuPheGlySerPheAlaAlaGlyArgThrSerAspTrpIleGlyArgArgTyrThr 106

Qy        366 ATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAAC 425
              ||||||   |||     |||||||||   |||      ||||||||||||||   |||
Db        107 IleValLeuAlaGlyAlaIlePhePheValGlyAlaLeuLeuMetGlyPheAlaThrAsn 126

Qy        426 TACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATG 485
              ||||||   ||||||   ||||||||||||   ||||||||||||||||||||||||||
Db        127 TyrAlaPheLeuMetValGlyArgPheValThrGlyIleGlyValGlyTyrAlaLeuMet 146

Qy        486 ATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCG 545
              ||||||||||||||||||||||||||||||||||||||:::|||||||||||||||||
Db        147 IleAlaProValTyrThrAlaGluValSerProAlaSerSerArgGlyPheLeuThrSer 166

Qy        546 TTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTC 605
              ||||||||||||||||||   ||||||||||||||||:::||||||   ||||||
Db        167 PheProGluValPheIleAsnAlaGlyIleLeuLeuGlyTyrIleSerAsnLeuAlaPhe 186

Qy        606 TCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCC 665
              |||   ||||||   :::|||   ||||||   ||||||||||||||||||   ||||||
Db        187 SerSerLeuProThrHisLeuSerTrpArgPheMetLeuGlyIleGlyAlaIleProSer 206

Qy        666 GTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAG 725
              :::   ||||||:::   |||||   ||||||||||||||||||||||||||||:::
Db        207 IlePheLeuAlaIleGlyValLeuAlaMetProGluSerProArgTrpLeuValMetGln 226

Qy        726 GGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCC 785
              ||||||||   ||||||||   ||||||   :::   |||||:::   ||||||||
Db        227 GlyArgLeuGlyAspAlaLysLysValLeuAsnArgIleSerAspSerProGluGluAla 246

Qy        786 GCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGAC 845
              ||||||::::::||||||      ||||||||||||   |||   |||   |||
Db        247 GlnLeuArgLeuSerGluIleLysGlnThrAlaGlyIleProAlaGluCysAspGluAsp 266

Qy        846 GTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATC 905
              :::      |||   |||          :::          ||||||||||||||
Db        267 IleTyrLysValGluLysThrLysIleLysSerGlyAsnAlaValTrpLysGluLeuPhe 286

Qy        906 CTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTC 965
                 :::||||||||||||:::||||||   :::::::::||||||||||||||||||||
Db        287 PheAsnProThrProAlaValArgArgAlaValIleAlaGlyIleGlyIleHisPhePhe 306

Qy        966 CAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCC 1025
              |||   |||   ||||||   :::||||||   ||||||||   :::|||:::|||
Db        307 GlnGlnAlaSerGlyIleAspAlaValValLeuTyrSerProArgIlePheGlnSerAla 326

Qy       1026 GGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGG 1085
              |||:::||||||   :::      |||      |||      .   ||||||   |||
Db        327 GlyIleThrAsnAlaArgLysGlnLeuLeuAlaThrValAlaValGlyValValLysThr 346
```

```
Qy       1086 CTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTG 1145
Db        347 LeuPheIleLeuValAlaThrPheGlnLeuAspLysTyrGlyArgArgProLeuLeuLeu 366

Qy       1146 GGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTC 1205
Db        367 ThrSerValGlyGlyMetIleIleAlaIleLeuThrLeuAlaMetSerLeuThrValIle 386

Qy       1206 GGCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTC 1265
Db        387 ---AspHisSerHisHisLysIleThrTrpAlaIleAlaLeuCysIleThrMetValCys 405

Qy       1266 GCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAG 1325
Db        406 AlaValValAlaSerPheSerIleGlyLeuGlyProIleThrTrpValTyrSerSerGlu 425

Qy       1326 ATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTC 1385
Db        426 ValPheProLeuArgLeuArgAlaGlnGlyThrSerMetGlyValAlaValAsnArgVal 445

Qy       1386 ACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGC 1445
Db        446 ValSerGlyValIleSerIlePhePheLeuProLeuSerHisLysIleThrThrGlyGly 465

Qy       1446 AGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTC 1505
Db        466 AlaPhePheLeuPheGlyGlyIleAlaIleIleAlaTrpPhePhePheLeuThrPheLeu 485

Qy       1506 CCGGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGAC--------- 1556
Db        486 ProGluThrArgGlyArgThrLeuGluAsnMetHisGluLeuPheGluAspPheArgTrp 505

Qy       1557 ------------ACGGCCGCCGCCTCGGAATCAGACGAGCCAGCC-------------- 1589
Db        506 ArgGluSerPheProGlyAsnLysSerAsnAsnAspGluAsnSerThrArgLysGlnSer 525

Qy       1590 ------AAGGAGAAGAAGAAGGTGGAAATGGCCGCCACTAACTGATCAAAC 1634
Db        526 AsnGlyAsnAspLysSerGlnValGlnLeuGlyGluThrThrThrSerSer 542

RESULT 9
P93075
ID   P93075      PRELIMINARY;      PRT;   549 AA.
AC   P93075;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   BvcDNA-205 protein.
GN   BVCDNA-205.
OS   Beta vulgaris (Sugar beet).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Caryophyllales; Amaranthaceae; Beta.
OX   NCBI_TaxID=161934;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97198558; PubMed=9046601;
RA   Chiou T.-J., Bush D.R.;
RT   "Isolation and molecular characteristics of two putative sugar
RT   transporters from sugar beet (Accession Nos. U64902 and U64903)
RT   (PGR97-017).";
RL   Plant Physiol. 113:663-663(1997).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR   EMBL; U64902; AAB68028.1; -.
DR   PIR; T14606; T14606.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0005351; F:sugar porter activity; IEA.
DR   GO; GO:0005215; F:transporter activity; IEA.
DR   GO; GO:0008643; P:carbohydrate transport; IEA.
DR   InterPro; IPR007114; MFS.
DR   InterPro; IPR005828; Sub_transporter.
DR   InterPro; IPR003663; Sugar_transpt.
DR   InterPro; IPR005829; Sug_transporter.
DR   Pfam; PF00083; sugar_tr; 1.
DR   PRINTS; PR00171; SUGRTRNSPORT.
DR   TIGRFAMs; TIGR00879; SP; 1.
DR   PROSITE; PS50850; MFS; 1.
DR   PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR   PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW   Sugar transport; Transmembrane; Transport.
SQ   SEQUENCE   549 AA;  59790 MW;  895615B299B43A69 CRC64;

Alignment Scores:
Pred. No.:              7.05e-85      Length:          549
Score:                  1589.00       Matches:         326
Percent Similarity:     72.23%        Conservative:    59
Best Local Similarity:  61.16%        Mismatches:      126
Query Match:            43.97%        Indels:          22
DB:                     10            Gaps:            5

US-10-051-902A-21 (1-2017) x P93075 (1-549)

Qy         78 TCCTCTGCACCACCGGAGATGGCTTCCGCCGCGCTG----------CCGGAGGCCGTC 125
Db         10 SerAspProProProThrThrAlaSerLysValIleAlaAspPheAspProLeuLysLys 29

Qy        126 GCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATG 185
Db         30 ProProLysArgAsn---------LysPheAlaPheAlaCysAlaThrLeuAlaSerMet 46

Qy        186 ACCTCCATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAG 245
Db         47 ThrSerValLeuLeuGlyTyrAspIleGlyValMetSerGlyAlaIleIleTyrLeuLys 66

Qy        246 AAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTAC 305
Db         67 GluAspTrpHisIleSerAspThrGlnIleGlyValLeuValGlyIleLeuAsnIleTyr 86

Qy        306 TCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACC 365
Db         87 CysLeuPheGlySerPheAlaAlaGlyArgThrSerAspTrpIleGlyArgArgTyrThr 106

Qy        366 ATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAAC 425
Db        107 IleValLeuAlaGlyAlaIlePhePheValGlyAlaLeuLeuMetGlyPheAlaThrAsn 126

Qy        426 TACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATG 485
Db        127 TyrAlaPheLeuMetValGlyArgPheValThrGlyIleGlyValGlyTyrAlaLeuMet 146

Qy        486 ATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCG 545
Db        147 IleAlaProValTyrThrAlaGluValSerProAlaSerSerArgGlyPheLeuThrSer 166

Qy        546 TTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTC 605
Db        167 PheProGluValPheIleAsnAlaGlyIleLeuLeuGlyTyrIleSerAsnLeuAlaPhe 186

Qy        606 TCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCC 665
Db        187 SerSerLeuProThrHisLeuSerTrpArgPheMetLeuGlyIleGlyAlaIleProSer 206

Qy        666 GTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAG 725
Db        207 IlePheLeuAlaIleGlyValLeuAlaMetProGluSerProArgTrpLeuValMetGln 226

Qy        726 GGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCC 785
Db        227 GlyArgLeuGlyAspAlaLysLysValLeuAsnArgIleSerAspSerProGluGluAla 246

Qy        786 GCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGAC 845
Db        247 GlnLeuArgLeuSerGluIleLysGlnThrAlaGlyIleProAlaGluCysAspGluAsp 266

Qy        846 GTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATC 905
```

```
Db        267 IleTyrLysValGluLysThrLysIleLysSerGlyAsnAlaValTrpLysGluLeuPhe 286

Qy        906 CTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTC 965
Db        287 PheAsnProThrProAlaValArgArgAlaValIleAlaGlyIleGlyIleHisPhePhe 306

Qy        966 CAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCC 1025
Db        307 GlnGlnAlaSerGlyIleAspAlaValValLeuTyrSerProArgIlePheGlnSerAla 326

Qy       1026 GGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGG 1085
Db        327 GlyIleThrAsnAlaArgLysGlnLeuLeuAlaThrValAlaValGlyValValLysThr 346

Qy       1086 CTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTG 1145
Db        347 LeuPheIleLeuValAlaThrPheGlnLeuAspLysTyrGlyArgArgProLeuLeuLeu 366

Qy       1146 GGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTC 1205
Db        367 ThrSerValGlyGlyMetIleIleAlaIleLeuThrLeuAlaMetSerLeuThrValIle 386

Qy       1206 GGCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTC 1265
Db        387 ---AspHisSerHisHisLysIleThrTrpAlaIleAlaLeuCysIleThrMetValCys 405

Qy       1266 GCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAG 1325
Db        406 AlaValValAlaSerPheSerIleGlyLeuGlyProIleThrTrpValTyrSerSerGlu 425

Qy       1326 ATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTC 1385
Db        426 ValPheProLeuArgLeuArgAlaGlnGlyThrSerMetGlyValAlaValAsnArgVal 445

Qy       1386 ACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGC 1445
Db        446 ValSerGlyValIleSerIlePhePheLeuProLeuSerHisLysIleThrThrGlyGly 465

Qy       1446 AGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTC 1505
Db        466 AlaPhePheLeuPheGlyGlyIleAlaIleIleAlaTrpPhePhePheLeuThrPheLeu 485

Qy       1506 CCGGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGAC--------- 1556
Db        486 ProGluThrArgGlyArgThrLeuGluAsnMetHisGluLeuPheGluAspPheArgTrp 505

Qy       1557 ------------ACGGCCGCCGCCTCGGAATCAGACGAGCCAGCC-------------- 1589
Db        506 ArgGluSerPheProGlyAsnLysSerAsnAsnAspGluAsnSerThrArgLysGlnSer 525

Qy       1590 ------AAGGAGAAGAAGAAGGTGGAAATGGCCGCCACT 1622
Db        526 AsnGlyAsnAspLysSerGlnValGlnLeuGlyGluThr 538
```

RESULT 10
Q9XIH7
ID Q9XIH7 PRELIMINARY; PRT; 511 AA.
AC Q9XIH7;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative sugar transporter.
GN AT2G16120.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana.";
RL Nature 402:761-768(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL; AC007134; AAD26954.1; -.
DR PIR; H84536; H84536.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0008643; P:carbohydrate transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00083; sugar_tr; 1.
DR PRINTS; PR00171; SUGRTRNSPORT.
DR TIGRFAMs; TIGR00879; SP; 1.
DR PROSITE; PS50850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Sugar transport; Transmembrane; Transport.
SQ SEQUENCE 511 AA; 54758 MW; E646A59316C6A54B CRC64;

Alignment Scores:

Pred. No.:	2.68e-84	Length:	511
Score:	1579.00	Matches:	306
Percent Similarity:	75.10%	Conservative:	77
Best Local Similarity:	60.00%	Mismatches:	121
Query Match:	43.69%	Indels:	6
DB:	10	Gaps:	2

US-10-051-902A-21 (1-2017) x Q9XIH7 (1-511)

```
Qy         96 ATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCG--------------CCGAAGAAGAAG 140
Db          1 MetAsnSerSerGlyValGluGlnGlyValValIleAlaGluSerGluProProArgGly 20

Qy        141 GGCAACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTC 200
Db         21 AsnArgSerArgTyrAlaPheAlaCysAlaIleLeuAlaSerMetThrSerIleIleLeu 40

Qy        201 GGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATC 260
Db         41 GlyTyrAspIleGlyValMetSerGlyAlaSerIlePheIleLysAspAspLeuLysLeu 60

Qy        261 AGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCC 320
Db         61 SerAspValGlnLeuGluIleLeuMetGlyIleLeuAsnIleTyrSerLeuValGlySer 80

Qy        321 TTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCC 380
Db         81 GlyAlaAlaGlyArgThrSerAspTrpLeuGlyArgArgTyrThrIleValLeuAlaGly 100

Qy        381 GTCATATTCTTCGCGGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATG 440
Db        101 AlaPhePhePheCysGlyAlaLeuLeuMetGlyPheAlaThrAsnTyrProPheIleMet 120

Qy        441 TTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTAC 500
Db        121 ValGlyArgPheValAlaGlyIleGlyValGlyTyrAlaMetMetIleAlaProValTyr 140
```

```
Qy        501 ACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTC 560
Db        141 ThrAlaGluValAlaProAlaSerSerArgGlyPheLeuThrSerPheProGluIlePhe 160

Qy        561 ATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTG 620
Db        161 IleAsnIleGlyIleLeuLeuGlyTyrValSerAsnTyrPhePheSerLysLeuProGlu 180

Qy        621 AACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTC 680
Db        181 HisLeuGlyTrpArgPheMetLeuGlyValGlyAlaValProSerValPheLeuAlaIle 200

Qy        681 ATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGAC 740
Db        201 GlyValLeuAlaMetProGluSerProArgTrpLeuValLeuGlnGlyArgLeuGlyAsp 220

Qy        741 GCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCC 800
Db        221 AlaPheLysValLeuAspLysThrSerAsnThrLysGluGluAlaIleSerArgLeuAsp 240

Qy        801 GACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCC 860
Db        241 AspIleLysArgAlaValGlyIleProAspAspMetThrAspAspValIleValValPro 260

Qy        861 AAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCG 920
Db        261 AsnLysLysSerAlaGly---LysGlyValTrpLysAspLeuLeuValArgProThrPro 279

Qy        921 GCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGC 980
Db        280 SerValArgHisIleLeuIleAlaCysLeuGlyIleHisPheAlaGlnGlnAlaSerGly 299

Qy        981 ATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGAC 1040
Db        300 IleAspAlaValValLeuTyrSerProThrIlePheSerLysAlaGlyLeuLysSerLys 319

Qy       1041 AAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTG 1100
Db        320 AsnAspGlnLeuLeuAlaThrValAlaValGlyValValLysThrLeuPheIleValVal 339

Qy       1101 GCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGG 1160
Db        340 GlyThrCysValValAspArgPheGlyArgArgAlaLeuLeuLeuThrSerMetGlyGly 359

Qy       1161 ATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGAC 1220
Db        360 MetPheLeuSerLeuThrAlaLeuGlyThrSerLeuThrValIleAsnArgAsnProGly 379

Qy       1221 GCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTC 1280
Db        380 GlnThrLeuLysTrpAlaIleGlyLeuAlaValThrThrValMetThrPheValAlaThr 399

Qy       1281 TTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAG 1340
Db        400 PheSerIleGlyAlaGlyProValThrTrpValTyrCysSerGluIlePheProValArg 419

Qy       1341 GTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATC 1400
Db        420 LeuArgAlaGlnGlyAlaSerLeuGlyValMetLeuAsnArgLeuMetSerGlyIleIle 439

Qy       1401 TCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTAC 1460
Db        440 GlyMetThrPheLeuSerLeuSerLysGlyLeuThrIleGlyGlyAlaPheLeuLeuPhe 459

Qy       1461 TCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGC 1520
Db        460 AlaGlyValAlaAlaAlaAlaTrpValPhePhePheThrPheLeuProGluThrArgGly 479

Qy       1521 CGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGAC 1580
Db        480 IleProLeuGluGluMetGluThrLeuPheGlySerTyrThrAlaAsnLysLysAsnAsn 499
```

```
Qy       1581 GAGCCAGCCAAGGAGAAGAAGAAGGTGGAA 1610
Db        500 SerMetSerLysAspAsnGluValValAsp 509

RESULT 11
Q9XIH6
ID   Q9XIH6      PRELIMINARY;      PRT;   511 AA.
AC   Q9XIH6;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Putative sugar transporter.
GN   AT2G16130.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=20083487; PubMed=10617197;
RA   Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA   Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA   Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA   Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA   Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA   Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA   Salzberg S.L., Fraser C.M., Venter J.C.;
RT   "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 402:761-768(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RA   Lin X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR   EMBL; AC007134; AAD26955.1; -.
DR   PIR; A84537; A84537.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0005351; F:sugar porter activity; IEA.
DR   GO; GO:0005215; F:transporter activity; IEA.
DR   GO; GO:0008643; P:carbohydrate transport; IEA.
DR   InterPro; IPR007114; MFS.
DR   InterPro; IPR005828; Sub_transporter.
DR   InterPro; IPR003663; Sugar_transpt.
DR   InterPro; IPR005829; Sug_transporter.
DR   Pfam; PF00083; sugar_tr; 1.
DR   PRINTS; PR00171; SUGRTRNSPORT.
DR   TIGRFAMs; TIGR00879; SP; 1.
DR   PROSITE; PS50850; MFS; 1.
DR   PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR   PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW   Sugar transport; Transmembrane; Transport.
SQ   SEQUENCE   511 AA;  54932 MW;  60CB43A9DB5C8396 CRC64;

Alignment Scores:
Pred. No.:                3.95e-83      Length:          511
Score:                    1559.00       Matches:         305
Percent Similarity:       75.15%        Conservative:    79
Best Local Similarity:    59.69%        Mismatches:      123
Query Match:              43.14%        Indels:          4
DB:                       10            Gaps:            3

US-10-051-902A-21 (1-2017) x Q9XIH6 (1-511)

Qy         81 TCTGCACCACCGGAGATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAG 140
Db          2 SerSerSerGlyGluGluArgGlyValValValAlaGluSerGluProPro------Arg 19

Qy        141 GGCAAC---GTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTC 197
```

```
Db        20 GlyAsnArgSerArgPheAlaPheAlaCysAlaIleLeuAlaSerMetThrSerIleIle 39

Qy       198 CTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAAC 257
Db        40 LeuGlyTyrAspIleGlyValMetSerGlyAlaAlaIlePheIleLysAspAspLeuLys 59

Qy       258 ATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGC 317
Db        60 LeuSerAspValGlnLeuGluIleLeuMetGlyIleLeuAsnIleTyrSerLeuIleGly 79

Qy       318 TCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCC 377
Db        80 SerGlyAlaAlaGlyArgThrSerAspTrpIleGlyArgArgTyrThrIleValLeuAla 99

Qy       378 GCCGTCATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTC 437
Db       100 GlyPhePhePhePheCysGlyAlaLeuLeuMetGlyPheAlaThrAsnTyrProPheIle 119

Qy       438 ATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTG 497
Db       120 MetValGlyArgPheValAlaGlyIleGlyValGlyTyrAlaMetMetIleAlaProVal 139

Qy       498 TACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTG 557
Db       140 TyrThrThrGluValAlaProAlaSerSerArgGlyPheLeuSerSerPheProGluIle 159

Qy       558 TTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCG 617
Db       160 PheIleAsnIleGlyIleLeuLeuGlyTyrValSerAsnTyrPhePheAlaLysLeuPro 179

Qy       618 CTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCG 677
Db       180 GluHisIleGlyTrpArgPheMetLeuGlyIleGlyAlaValProSerValPheLeuAla 199

Qy       678 CTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCG 737
Db       200 IleGlyValLeuAlaMetProGluSerProArgTrpLeuValMetGlnGlyArgLeuGly 219

Qy       738 GACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTG 797
Db       220 AspAlaPheLysValLeuAspLysThrSerAsnThrLysGluGluAlaIleSerArgLeu 239

Qy       798 GCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTC 857
Db       240 AsnAspIleLysArgAlaValGlyIleProAspAspMetThrAspAspValIleValVal 259

Qy       858 CCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACC 917
Db       260 ProAsnLysLysSerAlaGly---LysGlyValTrpLysAspLeuLeuValArgProThr 278

Qy       918 CCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTG 977
Db       279 ProSerValArgHisIleLeuIleAlaCysLeuGlyIleHisPheSerGlnGlnAlaSer 298

Qy       978 GGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAAC 1037
Db       299 GlyIleAspAlaValValLeuTyrSerProThrIlePheSerArgAlaGlyLeuLysSer 318

Qy      1038 GACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTG 1097
Db       319 LysAsnAspGlnLeuLeuAlaThrValAlaValGlyValValLysThrLeuPheIleVal 338

Qy      1098 TTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGC 1157
Db       339 ValGlyThrCysLeuValAspArgPheGlyArgArgAlaLeuLeuLeuThrSerMetGly 358

Qy      1158 GGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCC 1217
Db       359 GlyMetPhePheSerLeuThrAlaLeuGlyThrSerLeuThrValIleAspArgAsnPro 378

Qy      1218 GACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCC 1277

Db       379 GlyGlnThrLeuLysTrpAlaIleGlyLeuAlaValThrThrValMetThrPheValAla 398

Qy      1278 TTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTC 1337
Db       399 ThrPheSerLeuGlyAlaGlyProValThrTrpValTyrAlaSerGluIlePheProVal 418

Qy      1338 CAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTC 1397
Db       419 ArgLeuArgAlaGlnGlyAlaSerLeuGlyValMetLeuAsnArgLeuMetSerGlyIle 438

Qy      1398 ATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTC 1457
Db       439 IleGlyMetThrPheLeuSerLeuSerLysGlyLeuThrIleGlyGlyAlaPheLeuLeu 458

Qy      1458 TACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGC 1517
Db       459 PheAlaGlyValAlaValAlaAlaTrpValPhePhePheThrPheLeuProGluThrArg 478

Qy      1518 GGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCA 1577
Db       479 GlyValProLeuGluGluIleGluSerLeuPheGlySerTyrSerAlaAsnLysLysAsn 498

Qy      1578 GACGAGCCAGCCAAGGAGAAGAAGAAGGTGGAA 1610
Db       499 AsnValMetSerLysGlyLysGlnValValAsp 509
```

```
RESULT 12
Q84RI2
ID   Q84RI2      PRELIMINARY;      PRT;   491 AA.
AC   Q84RI2;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Sorbitol transporter.
GN   SOT1.
OS   Malus domestica (Apple) (Malus sylvestris).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids I; Rosales; Rosaceae; Maloideae; Malus.
OX   NCBI_TaxID=3750;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Mutsu; TISSUE=Fruit;
RA   Zhifang G., Jayanty S., Beaudry R.M., Loescher W.H.;
RT   "Cloning and characterization of apple fruit sorbitol transporter.";
RL   Submitted (FEB-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY237400; AAO88964.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005351; F:sugar porter activity; IEA.
DR   GO; GO:0005215; F:transporter activity; IEA.
DR   GO; GO:0008643; P:carbohydrate transport; IEA.
DR   InterPro; IPR007114; MFS.
DR   InterPro; IPR005828; Sub_transporter.
DR   InterPro; IPR003663; Sugar_transpt.
DR   InterPro; IPR005829; Sug_transporter.
DR   Pfam; PF00083; sugar_tr; 1.
DR   PRINTS; PR00171; SUGRTRNSPORT.
DR   TIGRFAMs; TIGR00879; SP; 1.
DR   PROSITE; PS50850; MFS; 1.
DR   PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR   PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
SQ   SEQUENCE   491 AA;  53310 MW;  3F9AE2E85CCB040A CRC64;
```

Alignment Scores:			
Pred. No.:	8.57e-81	Length:	491
Score:	1519.00	Matches:	288
Percent Similarity:	77.68%	Conservative:	81
Best Local Similarity:	60.63%	Mismatches:	104
Query Match:	42.03%	Indels:	2
DB:	10	Gaps:	1

US-10-051-902A-21 (1-2017) x Q84RI2 (1-491)

```
Qy       183 ATGACCTCCATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACATC 242
Db         1 MetThrSerIleLeuMetGlyTyrAspIleGlyValMetSerGlyAlaSerIleTyrIle 20

Qy       243 AAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTC 302
Db        21 GluLysAspLeuLysValThrAspThrGlnIleGluIleMetIleGlyValIleGluIle 40

Qy       303 TACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTAC 362
Db        41 TyrSerLeuIleGlySerAlaMetAlaGlyLysThrSerAspTrpValGlyArgArgTyr 60

Qy       363 ACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTC 422
Db        61 ThrIleValIleSerGlyAlaIlePhePheIleGlyAlaIleLeuMetGlyPheSerThr 80

Qy       423 AACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTC 482
Db        81 AsnTyrThrPheLeuMetCysGlyArgPheValAlaGlyIleGlyValGlyTyrAlaLeu 100

Qy       483 ATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACG 542
Db       101 ThrIleAlaProValTyrSerAlaGluValSerProThrSerSerArgGlyPheLeuThr 120

Qy       543 TCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCT 602
Db       121 SerPheProGluValPheValAsnIleGlyIleLeuLeuGlyTyrLeuSerAsnTyrAla 140

Qy       603 TTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCG 662
Db       141 PheSerPheCysProLeuAspLeuGlyTrpArgLeuMetLeuGlyValGlyAlaIlePro 160

Qy       663 TCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATG 722
Db       161 SerValGlyLeuAlaValGlyValLeuAlaMetProGluSerProArgTrpLeuValMet 180

Qy       723 AAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAG 782
Db       181 GlnGlyArgLeuGlyGluAlaLysArgValLeuAspArgThrSerAspSerLysGluGlu 200

Qy       783 GCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGC 842
Db       201 SerMetLeuArgLeuAlaAspIleLysGluAlaAlaGlyIleProGluGluCysAsnAsp 220

Qy       843 GACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTC 902
Db       221 AspIleValGlnVal------SerGlyHisSerHisGlyGluGlyValTrpLysGluLeu 238

Qy       903 ATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTC 962
Db       239 LeuValHisProThrProThrValArgHisIleLeuIleAlaAlaIleGlyPheHisPhe 258

Qy       963 TTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGC 1022
Db       259 PheGlnGlnAlaSerGlyIleAspAlaLeuValLeuTyrSerProArgValPheAlaLys 278

Qy      1023 CCCGGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAG 1082
Db       279 AlaGlyIleThrSerThrAsnGlnLeuLeuCysThrValGlyValGlyLeuSerLys 298

Qy      1083 AGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTG 1142
Db       299 ThrValPheThrLeuValAlaThrPhePheLeuAspArgValGlyArgArgProLeuLeu 318

Qy      1143 CTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTC 1202
Db       319 LeuThrSerMetAlaGlyMetValGlyAlaLeuValCysLeuGlyThrSerLeuThrIle 338

Qy      1203 GTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACC 1262
Db       339 ValAspGlnHisGluGlyValArgMetThrTrpAlaValIleLeuCysLeuCysCysVal 358

Qy      1263 CTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCG 1322
Db       359 LeuAlaTyrValGlyPhePheSerSerGlyIleGlyProIleAlaTrpValTyrSerSer 378

Qy      1323 GAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGC 1382
Db       379 GluIleProLeuArgLeuArgAlaGlnGlyCysGlyMetGlyValAlaValAsnArg 398

Qy      1383 GTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGC 1442
Db       399 LeuMetSerGlyIleLeuSerMetThrPheIleSerLeuTyrLysAlaIleThrMetGly 418

Qy      1443 GGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTAC 1502
Db       419 GlyThrPhePheLeuTyrAlaAlaIleGlyThrValGlyTrpIlePhePhePheThrMet 438

Qy      1503 CTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGGCC 1562
Db       439 LeuProGluThrGlnGlyArgThrLeuGluAspMetGluValLeuPheGlyLysPheHis 458

Qy      1563 GCCGCCTCGGAATCAGACGAGCCAGCCAAGGAGAAGAAGAAGGTG 1607
Db       459 LysTrpArgLysAlaAsnLysLeuLeuGluLysLysLysArgVal 473
```

RESULT 13
Q9FQX3
ID Q9FQX3 PRELIMINARY; PRT; 513 AA.
AC Q9FQX3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Mannitol transporter.
GN MAT1.
OS Apium graveolens var. dulce.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC campanulids; Apiales; Apiaceae; Apioideae; apioid superclade;
OC Apium clade; Apium.
OX NCBI_TaxID=117781;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Phloem;
RX MEDLINE=21149881; PubMed=11251106;
RA Noiraud N., Maurousset L., Lemoine R.;
RT "Identification of a mannitol transporter, agmat1, in celery phloem.";
RL Plant Cell 13:695-705(2001).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL; AF215837; AAG43998.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0008643; P:carbohydrate transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00083; sugar_tr; 1.
DR PRINTS; PR00171; SUGRTRNSPORT.
DR TIGRFAMs; TIGR00879; SP; 1.
DR PROSITE; PS50850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Sugar transport; Transmembrane; Transport.
SQ SEQUENCE 513 AA; 56013 MW; E08D6555B9E08781 CRC64;

Alignment Scores:

Pred. No.:	1.94e-80	Length:	513
Score:	1513.00	Matches:	295
Percent Similarity:	76.77%	Conservative:	85
Best Local Similarity:	59.60%	Mismatches:	109
Query Match:	41.86%	Indels:	6
DB:	10	Gaps:	4

```
US-10-051-902A-21 (1-2017) x Q9FQX3 (1-513)

Qy        129 CCGAAGAAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACC 188
              ||||||   |||   |||   :::::::|||||||||||||||||:::|||||||||||
Db         16 ProLysProLysArgAsn---LysTyrAlaPheAlaCysAlaLeuLeuAlaSerMetAsn 34

Qy        189 TCCATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAG 248
              ||||||||||||||||||||   ||||||:::|||||||||||||:::||||||||:::
Db         35 SerIleLeuLeuGlyTyrAspThrGlyValLeuSerGlyAlaSerIleTyrIleLysGlu 54

Qy        249 GACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCG 308
              |||   :::   ||||||   :::::::|||:::::::::|||||||:::|||:::||||||
Db         55 AspLeuHisPheSerAspValGlnIleGluIleIleIleGlyIleIleAsnIleTyrSer 74

Qy        309 CTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATC 368
              |||:::||||||      |||||||||||||||||||||||||||||||||||||:::
Db         75 LeuLeuGlySerAlaIleAlaGlyArgThrSerAspTrpIleGlyArgArgTyrThrMet 94

Qy        369 GTGTTCGCCGCCGTCATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTAC 428
              |||   |||   :::|||||||||   |||          ||||||   |||   |||:::
Db         95 ValLeuAlaGlyIleIlePhePheLeuGlyAlaIlePheMetGlyLeuAlaThrAsnPhe 114

Qy        429 GCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATC 488
              |||   ||||||||||||||||||||||||||||||||||||||||||||:::||||||
Db        115 AlaPheLeuMetPheGlyArgPheValAlaGlyIleGlyValGlyTyrAlaMetMetIle 134

Qy        489 GCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTC 548
              |||||||||||||||||||||||:::|||:::|||:::||||||||||||||||||||
Db        135 AlaProValTyrThrAlaGluValAlaProSerSerSerArgGlyPheLeuThrSerPhe 154

Qy        549 CCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCC 608
              ||||||||||||||||   |||:::|||||||||||||||||||||:::||||||:::
Db        155 ProGluValPheIleAsnSerGlyValLeuLeuGlyTyrValSerAsnPheAlaPheAla 174

Qy        609 CGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTG 668
              :::   ||||||   |||||||||||||||||||||||||||||||||   |||||||||
Db        175 LysCysProLeuTrpLeuGlyTrpArgIleMetLeuGlyIleGlyAlaPheProSerVal 194

Qy        669 CTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGA 728
              ||||||:::::||||||   ||||||||||||||||||||||||||||:::|||
Db        195 AlaLeuAlaIleIleValLeuTyrMetProGluSerProArgTrpLeuValMetGlnGly 214

Qy        729 CGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCG 788
              ||||||   :::|||:::   |||||||||||||||||   :::   |||||||||
Db        215 ArgLeuGlyGluAlaArgThrValLeuGluLysThrSerThrSerLysGluGluAlaHis 234

Qy        789 GAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTG 848
              :::|||||:::|||||||||   ||||||||||||   ::::::   :::   ||||||
Db        235 GlnArgLeuSerAspIleLysGluAlaAlaGlyIleAspLysAspCysAsnAspAspVal 254

Qy        849 GTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTG 908
              |||   ||||||||||||           ::::::   ||||||||||||||||||||
Db        255 ValGlnValProLysArg------ThrLysAspGluAlaValTrpLysGluLeuIleLeu 272

Qy        909 TCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAG 968
              ||||||      :::|||         ::::::||||||||||||||||||||||||
Db        273 HisProThrLysProValArgHisAlaAlaIleThrGlyIleGlyIleHisPhePheGln 292

Qy        969 CATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGA 1028
              |||   ||||||   :::|||||||   ||||||||   :::|||:::      |||
Db        293 GlnAlaCysGlyIleAspAlaValValLeuTyrSerProArgIlePheGluLysAlaGly 312

Qy       1029 TTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTT 1088
              :::   ::::::          |||       |||          ||||||   |||   :::
Db        313 IleLysSerAsnSerLysLysLeuLeuAlaThrIleAlaValGlyValCysLysThrVal 332

Qy       1089 TTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGC 1148
              |||||||||::::::||||||   :::|||   :::|||||||||||||||:::|||
Db        333 PheIleLeuIleSerThrPheGlnLeuAspLysIleGlyArgArgProLeuMetLeuThr 352

Qy       1149 AGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGC 1208
              |||   |||||:::::::::::|||      |||          |||||||||:::
Db        353 SerMetGlyGlyMetValIleAlaLeuPheValLeuAlaGlySerLeuThrValIleAsn 372

Qy       1209 CAG---CACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTC 1265
                :::   |||                    ||||||   ||||||:::|||   :::
Db        373 LysSerHisHisThrGlyHis------TrpAlaGlyGlyLeuAlaIlePheThrValTyr 390

Qy       1266 GCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAG 1325
              |||:::|||:::   ||||||   |||:::|||||||||   ||||||||||||||||||
Db        391 AlaPheValSerIlePheSerSerGlyMetGlyProIleAlaTrpValTyrSerSerGlu 410

Qy       1326 ATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTC 1385
              :::|||||||||:::::|||||   ||||||||:::|||||||   |||||||
Db        411 ValPheProLeuArgLeuArgAlaGlnGlyCysSerIleGlyValAlaValAsnArgGly 430

Qy       1386 ACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGC 1445
                ||||||:::|||   |||||||||:::|||:::   ||||||:::||||||||||||
Db        431 MetSerGlyIleIleGlyMetThrPheIleSerMetTyrLysAlaMetThrIleGlyGly 450

Qy       1446 AGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTC 1505
              :::|||   |||:::::   :::|||:::::   |||||||||   ||||||
Db        451 AlaPheLeuLeuPheAlaValValAlaSerIleGlyTrpValPheMetTyrThrMetPhe 470

Qy       1506 CCGGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGGCCGCC 1565
              |||||||||:::||||||   |||||||||:::      ||||||||
Db        471 ProGluThrGlnGlyArgAsnLeuGluGluIleGluLeuLeuPheGlySerTyrPheGly 490

Qy       1566 GCCTCGGAATCAGACGAGCCAGCCAAGGAGAAGAAGAAGGTGGAA 1610
                  ::::::   :::      |||   |||:::      |||
Db        491 TrpArgLysThrLeuLysAspLeuLysAlaLysGluAlaAlaGlu 505

RESULT 14
Q7XB36
ID   Q7XB36      PRELIMINARY;      PRT;   519 AA.
AC   Q7XB36;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Putative mannitol transporter.
GN   MAT1.
OS   Orobanche ramosa.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   lamiids; Lamiales; Orobanchaceae; Orobancheae; Orobanche.
OX   NCBI_TaxID=46066;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Delavault P., Simier P., Le Coguic O.;
RT   "Identification of a mannitol transporter, OrMaT1, in Orobanche
RT   ramosa.";
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY136668; AAN07021.1; -.
SQ   SEQUENCE   519 AA;  56257 MW;  97A54D8D0F20047C CRC64;

Alignment Scores:
Pred. No.:              2.51e-79        Length:         519
Score:                  1494.00         Matches:        289
Percent Similarity:     77.13%          Conservative:   82
Best Local Similarity:  60.08%          Mismatches:     104
Query Match:            41.34%          Indels:         6
DB:                     10              Gaps:           3

US-10-051-902A-21 (1-2017) x Q7XB36 (1-519)

Qy        117 GAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGCGCCATCCTC 176
              :::|||:::      |||   |||   |||:::   :::|||         :::::::::
Db         20 LysAlaLeuIleAsnLysProLysLysAsnMet---TyrAlaLeuLeuIleSerValMet 38

Qy        177 GCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTG 236
              ||||||||||||||||||||||||||||||   ||||||||||||||||:::|||
Db         39 AlaSerMetThrSerIleLeuLeuGlyTyrAspThrGlyValMetSerGlyAlaThrLeu 58

Qy        237 TACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTG 296
              ||||||||||||||      |||:::|||   :::||||||:::|||:::|||::::::
Db         59 TyrIleLysLysAspLeuLysIleThrAspValGlnValGluIleLeuValGlyLeuIle 78
```

```
Qy        297 AACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGGCGGACGTCGGACTGGATCGGCCGG 356
Db         79 AsnIleTyrSerLeuIleGlySerAlaValAlaGlyArgThrSerAspTyrLeuGlyArg 98

Qy        357 CGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGGSGTTCCTCATGGGGTTC 416
Db         99 ArgIleThrIleValIleAlaSerValIlePhePheValGlyAlaAlaValMetGlyLeu 118

Qy        417 GCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTAC 476
Db        119 AlaAsnAsnTyrAlaValLeuMetValGlyArgPheValAlaGlyLeuGlyValGlyTyr 138

Qy        477 GCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTC 536
Db        139 AlaLeuMetIleAlaProValTyrAlaAlaGluValAlaProAlaSerSerArgGlyPhe 158

Qy        537 CTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAAC 596
Db        159 LeuThrSerPheProGluValPheIleAsnPheGlyValLeuLeuGlyTyrLeuSerAsn 178

Qy        597 TATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCG 656
Db        179 TyrAlaPheAlaLysPheSerLeuLysLeuGlyTrpArgLeuMetLeuGlyValGlyAla 198

Qy        657 GCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTG 716
Db        199 LeuProAlaIlePheIleGlyLeuAlaValIleValMetProGluSerProArgTrpLeu 218

Qy        717 GTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCG 776
Db        219 ValMetGlnGlyArgLeuGlyAspAlaLysLysValLeuAspArgThrSerAspSerPro 238

Qy        777 GAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTC 836
Db        239 GlnGluAlaGlnLeuArgLeuAlaAspIleMetGluAlaAlaGlyLeuProGluAspCys 258

Qy        837 GACGGCGACGTGGTGACCGTCCCCAAG------AGAGGGAGCGGAAACGAGAAGCGGGTG 890
Db        259 HisAspAspValValProValLeuLysGlnAspArgGlyGlyGlyGly---------Val 275

Qy        891 TGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATC 950
Db        276 TrpLysGluLeuIleValHisProThrLysProValLeuHisIleThrIleAlaAlaVal 295

Qy        951 GGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTC 1010
Db        296 GlyCysGlnPhePheGlnGlnAlaSerGlyIleAspAlaValValMetTyrSerProArg 315

Qy       1011 GTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTC 1070
Db        316 IleTyrGluLysAlaGlyIleThrSerAspGluLysLysLeuLeuAlaThrIleAlaVal 335

Qy       1071 GGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGG 1130
Db        336 GlyLeuCysLysThrValPheIleLeuValThrThrPheMetValAspArgIleGlyArg 355

Qy       1131 CGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCC 1190
Db        356 ArgValLeuLeuLeuThrSerCysGlyGlyLeuValLeuSerMetLeuThrLeuAlaThr 375

Qy       1191 GGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGC 1250
Db        376 GlyLeuThrValIleAspHisTyrGlyAlaAspArgPheProTrpValValValLeuCys 395

Qy       1251 ATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGG 1310
Db        396 ValLeuThrThrTyrSerSerValAlaPhePheSerMetGlyMetGlyProIleAlaTrp 415

Qy       1311 GTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTC 1370
Db        416 ValTyrSerSerGluIlePheProLeuLysLeuArgAlaGlnGlyCysGlyLeuGlyVal 435

Qy       1371 GCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCC 1430
Db        436 AlaIleAsnArgAlaThrAsnGlyValIleLeuMetSerPheIleSerLeuTyrAsnAla 455

Qy       1431 ATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTC 1490
Db        456 IleThrIleGlyGlyAlaPheTyrLeuPheSerGlyIleGlyIleValThrTrpIlePhe 475

Qy       1491 TTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTC 1550
Db        476 PhePheThrLeuLeuProGluThrArgGlyArgThrLeuGluAspMetGluValLeuPhe 495

Qy       1551 GGC 1553
Db        496 Gly 496
```

RESULT 15
Q8RVQ2
ID   Q8RVQ2   PRELIMINARY;   PRT;   523 AA.
AC   Q8RVQ2;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Mannitol transporter.
GN   MAT2.
OS   Apium graveolens var. dulce.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   campanulids; Apiales; Apiaceae; Apioideae; apioid superclade;
OC   Apium clade; Apium.
OX   NCBI_TaxID=117781;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Vert d'Elne; TISSUE=Petiole phloem;
RA   Noiraud N., Pichaud J.-P., Maurousset L., Lemoine R.;
RT   "AgMaT2 a mannitol transporter from celery.";
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR   EMBL; AF480069; AAL85876.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0005351; F:sugar porter activity; IEA.
DR   GO; GO:0005215; F:transporter activity; IEA.
DR   GO; GO:0008643; P:carbohydrate transport; IEA.
DR   InterPro; IPR007114; MFS.
DR   InterPro; IPR005828; Sub_transporter.
DR   InterPro; IPR003663; Sugar_transpt.
DR   InterPro; IPR005829; Sug_transporter.
DR   Pfam; PF00083; sugar_tr; 1.
DR   TIGRFAMs; TIGR00879; SP; 1.
DR   PROSITE; PS50850; MFS; 1.
DR   PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
KW   Sugar transport; Transmembrane; Transport.
SQ   SEQUENCE   523 AA; 56754 MW; C01D0602BB3D3575 CRC64;

Alignment Scores:

Pred. No.:	1.03e-73	Length:	523
Score:	1398.00	Matches:	286
Percent Similarity:	70.60%	Conservative:	79
Best Local Similarity:	55.32%	Mismatches:	117
Query Match:	38.68%	Indels:	36
DB:	10	Gaps:	6

US-10-051-902A-21 (1-2017) x Q8RVQ2 (1-523)

```
Qy         42 GCCACTACTGTACACGGCCCAGAGCGAGCCTCCTCCTCCTCTGCACCACCGGAGATGGCT 101
Db          4 AlaPheAsnValSerGlyAspAspSerGlyGlySerLys----------AspValGly 19

Qy        102 TCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCCTTC 161
Db         20 SerProValIlePheAspAlaVal------LysLysProLysArgAsnLysTyrAlaPhe 37
```

```
Qy      162 GCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTGATG 221
            |||||||||||||||||||||||||||:::|||||||||||||||||||||||||||||
Db       38 GlyCysAlaIleLeuAlaSerMetThrSerValLeuLeuGlyTyrAspIleGlyValMet 57

Qy      222 AGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTT 281
            ||||||||:::::|||||||||        ::::::||||||   |||:::|||:::
Db       58 SerGlyAlaAlaIleTyrIleLysAspGlnLeuHisValSerAspValLysLeuGluIle 77

Qy      282 CTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCG 341
            ::::::|||||:::|||   :::||||||:::||||||      ||||||||||||||||
Db       78 ValValGlyIleIleAsnPhePheSerLeuValGlySerAlaLeuAlaGlyArgThrSer 97

Qy      342 GACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGSG 401
            |||||||||||||||||||||||:::|||   |||      |||||||||   |||
Db       98 AspTrpIleGlyArgArgTyrThrMetValLeuAlaGlyAlaIlePhePheValGlyAla 117

Qy      402 TTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGC 461
            ||||||||||||||   ||||||:::   ||||||||||||||||||||||||||||||
Db      118 IleLeuMetGlyPheAlaThrAsnTyrSerPheLeuMetPheGlyArgPheValAlaGly 137

Qy      462 AT-CGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGC 520
            |||      |||      ||||||   :::|||||||||   ||||||
Db      138 IleArgSerArgLeuCysProHisAspCysSerGlyValHisGlyArgGlyPhePheSer 157

Qy      521 GTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCT 580
            :::   :::|||:::         :::||||||    |||||||||   |||:::|||||
Db      158 IleIleSerTrpIleSerTyrPheIleProGlySer-PheIleAsnIleGlyValLeuLe 177

Qy      581 CGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCAT 640
            |||||||||||||||||||||||||||:::||||||   ||||||||||||||   ||
Db      177 uGlyTyrValSerAsnTyrAlaPheSerLysLeuProAlaAsnLeuGlyTrpArgPheMe 197

Qy      641 GCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGA 700
            ||||||||||||||||   ||||||:::   ||||||:::   ||||||||||||||||
Db      197 tLeuGlyIleGlyAlaIleProSerIleGlyLeuAlaIleGlyValLeuGlyMetProGl 217

Qy      701 GTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAA 760
            |||||||||||||||||||||||||||||||||   :::|||:::   ||||||:::||
Db      217 uSerProArgTrpLeuValMetLysGlyArgLeuGlyGluAlaArgGlnValLeuAspLy 237

Qy      761 GACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGG 820
            ||||||||||:::   ||||||:::      |||||:::||||||||   ||||||||
Db      237 sThrSerAspSerLysGluGluSerArgLeuArgLeuSerAspIleLysGlnAlaAlaGl 257

Qy      821 CATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGA 880
            ||||||||||||   :::   |||:::|||   :::||||||||||          |||::
Db      257 yIleProGluGluCysAsnAspAspIleValValMetProLysArg------ArgAsnAs 275

Qy      881 GAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCT 940
            ::::   |||||||||||||||:::|||   ||||||||:::::|||         ::
Db      275 pGluAlaValTrpLysGluLeuLeuLeuHisProThrProSerValArgHisAlaPheIl 295

Qy      941 GTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTA 1000
            ::::|||:::|||:::||||||||||||   :::   ||||||   :::      ||
Db      295 eAlaGlyValGlyLeuHisPhePheGlnGlnSerSerGlyIleAspAlaGlyGlyLeuTy 315

Qy     1001 CAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGCACCAC 1060
            |||||||   :::|||:::      |||:::|||:::             |||      ||
Db      315 rSerProArgIlePheGluLysAlaGlyIleThrSerThrAspLeuLysLeuLeuAlaTh 335

Qy     1061 TTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGG 1120
            |         |||::::::|||   ||||||||||||:::|||||||||   :::|||
Db      335 rIleAlaValGlyIleSerLysThrLeuPheIleLeuValAlaThrPheLeuLeuAspAr 355

Qy     1121 CGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATCGG 1180
             :::|||||||||||||||||||   |||   ||||||:::|||:::||||||
Db      355 gIleGlyArgArgProLeuLeuLeuThrSerMetGlyGlyMetIleIleSerLeuThrLe 375

Qy     1181 CCTCGGCGCCGGGCTCACCGTCGTCGGC------------------------------ 1208
             ||||||      |||   |||:::
Db      375 uLeuGlyThrSerLeuAlaValIleAlaThrGlnIleThrGlnPheIleGlyProGlyHi 395

Qy     1209 ----CAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCT 1264
```

```
                |||      :::           ||||||:::
Db      395 sTrpGlnSerSerGluValTyrHisThrTrpAlaVal---------------------- 407

Qy     1265 CGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGA 1324
                       ||||||      ||||||||||||   |||   ||||||||||
Db      408 ------------PhePhe------TrpLeuGlyProIleAlaTrpGlyTyrSerSerGl 421

Qy     1325 GATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGT 1384
            |:::|||||||||:::::||||||   |||||||||:::||||||||   ||||||
Db      421 uValPheProLeuArgLeuArgAlaGlnGlyCysSerIleGlyValAlaValAsnArgGl 441

Qy     1385 CACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGG 1444
             |||||||||:::|||||||||||||||||||||||   |||||||||:::|||   ||
Db      441 yThrSerGlyIleIleSerMetThrPheLeuSerLeuTyrLysAlaIleSerIleAlaGl 461

Qy     1445 CAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCT 1504
            |:::|||:::|||:::::   ||||||   :::||||||:::|||   :::|||   ||
Db      461 yAlaPheTyrLeuPheAlaAlaIleAlaGlyValAlaTrpIlePheIlePheThrLeuLe 481

Qy     1505 CCCGGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGC 1553
            ||||||||||:::||||||:::||||||||||||      |||||||||
Db      481 uProGluThrGlnGlyArgSerLeuGluGluMetGlyLeuLeuPheGly 497
```

Search completed: June 30, 2004, 18:52:02
Job time : 154.5 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: July 2, 2004, 02:11:44 ; Search time 12597 Seconds
(without alignments)
6939.978 Million cell updates/sec

Title: US-10-051-902A-21
Perfect score: 2017
Sequence: 1 cttacatgtaagctcgtgcc.....aaaaaaaaaaaaaaaaaaaa 2017

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		2016.6	100.0	2017	6	AR208575	AR208575 Sequence
2		1887	93.6	1961	8	AK102676	AK102676 Oryza sat
3		1880.6	93.2	1962	8	AK059423	AK059423 Oryza sat
4		1783	88.4	25270	8	AP004185	AP004185 Oryza sat
5		1660.6	82.3	1860	8	AF416867	AF416867 Oryza sat
6		1120.2	55.5	2109	8	BT009264	BT009264 Triticum
7		1094.6	54.3	2089	6	AR208577	AR208577 Sequence
8		1032.2	51.2	1620	8	AK119464	AK119464 Oryza sat
9		1005	49.8	1914	6	AR208574	AR208574 Sequence
10		988	49.0	1872	8	BT009301	BT009301 Triticum
11	c	985.8	48.9	25270	8	AP004185	AP004185 Oryza sat
12		981.6	48.7	1872	6	AR208578	AR208578 Sequence
13	c	909.2	45.1	193577	8	AP004365	AP004365 Oryza sat
14	c	759.4	37.6	155037	8	AC116426	AC116426 Genomic s
15	c	759.4	37.6	181163	8	AC115687	AC115687 Genomic s
16		739.6	36.7	2300	8	AK070417	AK070417 Oryza sat
17		738	36.6	2167	8	AK103915	AK103915 Oryza sat
18		729.2	36.2	2252	8	AK099273	AK099273 Oryza sat
19		687.8	34.1	139123	8	AC145480	AC145480 Genomic s
20		657.4	32.6	1998	8	AK121279	AK121279 Oryza sat
21		644.4	31.9	1927	8	AF482011	AF482011 Prunus ce
22		628.6	31.2	2032	8	AY100638	AY100638 Prunus ce
23		620	30.7	141050	8	AC093093	AC093093 Oryza sat
24		620	30.7	147706	8	AC083945	AC083945 Oryza sat
25	c	620	30.7	301337	8	AE017082	AE017082 Oryza sat
26		603.2	29.9	1648	8	AY136668	AY136668 Orobanche
27		600.4	29.8	2012	8	AY237400	AY237400 Malus x d
28		590.2	29.3	1729	8	AK061613	AK061613 Oryza sat
29		590.2	29.3	1857	8	AK099395	AK099395 Oryza sat
30		588	29.2	1784	8	AY237401	AY237401 Malus x d
31		576.2	28.6	1987	8	PMA532589	AJ532589 Plantago
32		575.2	28.5	1853	6	AR208576	AR208576 Sequence
33		573.6	28.4	1614	8	AK106745	AK106745 Oryza sat
34		561.2	27.8	1894	8	GMA563367	AJ563367 Glycine m
35		548.8	27.2	1638	8	AK119762	AK119762 Oryza sat
36		546.8	27.1	133331	8	OSJN00017	AL606456 Oryza sat
37		545.2	27.0	103264	2	OSIG00049	AL732353 Oryza sat
38		512.6	25.4	1946	8	AK121039	AK121039 Oryza sat
39		510.8	25.3	1891	8	PMA532590	AJ532590 Plantago
40	c	504.6	25.0	162766	8	AC125782	AC125782 Oryza sat
41		504.6	25.0	171946	8	AC145349	AC145349 Oryza sat
42	c	504.6	25.0	348970	2	AC138608	AC138608 Oryza sat
43		477.2	23.7	2054	8	BVU64903	U64903 Beta vulgar
44		477.2	23.7	2056	6	AX353437	AX353437 Sequence
45		469.2	23.3	1925	6	AX353438	AX353438 Sequence

ALIGNMENTS

RESULT 1
AR208575
LOCUS AR208575 2017 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 21 from patent US 6383776.
ACCESSION AR208575
VERSION AR208575.1 GI:21509765
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 2017)
AUTHORS Allen,S.M., Hitz,W.D., Kinney,A.J. and Tingey,S.V.
TITLE Plant sugar transport proteins
JOURNAL Patent: US 6383776-A 21 07-MAY-2002;
FEATURES Location/Qualifiers

source 1. .2017
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN

Query Match 100.0%; Score 2016.6; DB 6; Length 2017;
Best Local Similarity 100.0%; Pred. No. 4e-265;
Matches 2017; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CTTACATGTAAGCTCGTGCCGGCACGAGCTTACACTCGACCGCCACTACTGTACACGGCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTTACATGTAAGCTCGTGCCGGCACGAGCTTACACTCGACCGCCACTACTGTACACGGCC 60

Qy         61 CAGAGCGAGCCTCCTCCTCCTCTGCACCACCGGAGATGGCTTCCGCCGCGCTGCCGGAGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGAGCGAGCCTCCTCCTCCTCTGCACCACCGGAGATGGCTTCCGCCGCGCTGCCGGAGG 120

Qy        121 CCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCT 180

Qy        181 CCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACA 240

Qy        241 TCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACC 300

Qy        301 TCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGT 360

Qy        361 ACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCG 420

Qy        421 TCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGC 480

Qy        481 TCATGATCGCGCCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCATGATCGCGCCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGA 540

Qy        541 CGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATG 600

Qy        601 CTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGC 660

Qy        661 CGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCA 720

Qy        721 TGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGG 780

Qy        781 AGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACG 840

Qy        841 GCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGC 900

Qy        901 TCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACT 960

Qy        961 TCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGA 1020

Qy       1021 GCCCCGGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GCCCCGGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCA 1080

Qy       1081 AGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGT 1140

Qy       1141 TGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCG 1200

Qy       1201 TCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCA 1260

Qy       1261 CCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCT 1320

Qy       1321 CGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACC 1380

Qy       1381 GCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 GCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCG 1440

Qy       1441 GCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCT 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCT 1500

Qy       1501 ACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 ACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGG 1560

Qy       1561 CCGCCGCCTCGGAATCAGACGAGCCAGCCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCA 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CCGCCGCCTCGGAATCAGACGAGCCAGCCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCA 1620

Qy       1621 CTAACTGATCAAACTAACCGCAAAATCACCAAATCCTAAGGGTTTTCTTGCAAAAACGTG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 CTAACTGATCAAACTAACCGCAAAATCACCAAATCCTAAGGGTTTTCTTGCAAAAACGTG 1680

Qy       1681 TGCTGTACTGGCTAGCTAGCAAGTAGTAGCAGCAACGTGGGAAGATTCGCTGATCCGGCG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 TGCTGTACTGGCTAGCTAGCAAGTAGTAGCAGCAACGTGGGAAGATTCGCTGATCCGGCG 1740

Qy       1741 TTGCTGGAGAGCGACGGCCGGCGACGACAAAGCTGAGCTCCAGCTCGAGACTTCTTAAAA 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TTGCTGGAGAGCGACGGCCGGCGACGACAAAGCTGAGCTCCAGCTCGAGACTTCTTAAAA 1800

Qy       1801 TCATCTTCAAGTACATGGATTTTATTTTGCTCTTTGCTTTGTCCGTAAAAGTTGTACTAT 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 TCATCTTCAAGTACATGGATTTTATTTTGCTCTTTGCTTTGTCCGTAAAAGTTGTACTAT- 1860

Qy       1861 GCGATGAAGAATACCAGTATGTAGCAAGGCTGAGGTTGTGTGTAGCTACTAGAAGTGTCA 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 GCGATGAAGAATACCAGTATGTAGCAAGGCTGAGGTTGTGTGTAGCTACTAGAAGTGTCA 1920

Qy       1921 GTCACGTTGTTCTTGTAAGAAATGTTTAACTGTTAATTAAGCAGTATTGTTGCAGTAATC 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GTCACGTTGTTCTTGTAAGAAATGTTTAACTGTTAATTAAGCAGTATTGTTGCAGTAATC 1980

Qy       1981 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2017
              |||||||||||||||||||||||||||||||||||||
Db       1981 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2017
```

RESULT 2
AK102676
LOCUS AK102676 1961 bp mRNA linear PLN 24-JUL-2003
DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:J033102G08, full insert sequence.
ACCESSION AK102676
VERSION AK102676.1 GI:32987885
KEYWORDS FLI_CDNA; CAP trapper.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1
AUTHORS The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team:, Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group:, Otomo,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y., Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M., Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J., Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J., Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN:, Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y., Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Yoshino,M. and Hayashizaki,Y.
TITLE Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice
JOURNAL Science 301 (5631), 376-379 (2003)
MEDLINE 22752273
PUBMED 12869764
REFERENCE 2 (bases 1 to 1961)
AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K., Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K., Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M., Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M., Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A., Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M., Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K., Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H., Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K., Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S., Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W., Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and Yoshimura,A.
TITLE Direct Submission
JOURNAL Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)
COMMENT This clone is one of the 28K full-length cDNA clones from japonica rice.
URL : http://cdna01.dna.affrc.go.jp/cDNA/
NIAS Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and Yamamoto,M.
FAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y., Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M., Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J., Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S., Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S., Yoshimura,A., Matsubara,K. and Murakami,K.
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J., Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N., Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K., Yasunishi,A. and Hayashizaki,Y.

FEATURES Location/Qualifiers
source 1. .1961
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J033102G08"

ORIGIN

Query Match 93.6%; Score 1887; DB 8; Length 1961;
Best Local Similarity 97.8%; Pred. No. 1.8e-247;
Matches 1911; Conservative 1; Mismatches 41; Indels 0; Gaps 0;

```
Qy         28 GCTTACACTCGACCGCCACTACTGTACACGGCCCAGAGCGAGCCTCCTCCTCCTCTGCAC 87
              ||    |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          9 GCGAGCTCTCGACCGCCACTACTGTACACGGCCCAGAGCGAGCCTCCTCCTCCTCTGCAC 68

Qy         88 CACCGGAGATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACG 147
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         69 CACCGGAGATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACG 128

Qy        148 TCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACG 207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        129 TCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACG 188

Qy        208 ATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACG 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        189 ATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACG 248

Qy        268 GGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGG 327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 GGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGG 308

Qy        328 CGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATAT 387
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 CGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATAT 368

Qy        388 TCTTCGCGGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCC 447
              ||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        369 TCTTCGCGGGGGCGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCC 428

Qy        448 GCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCG 507
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        429 GCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCG 488

Qy        508 AGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACT 567
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        489 AGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACT 548

Qy        568 TCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCG 627
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        549 TCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCG 608

Qy        628 GGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGC 687
```

```
Db        609 GGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGC 668
Qy        688 TCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGG 747
Db        669 TCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGG 728
Qy        748 TGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCA 807
Db        729 TGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCA 788
Qy        808 AGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAG 867
Db        789 AGGCCGCCGCCGGCATCCCCGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAG 848
Qy        868 GGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGC 927
Db        849 GGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGC 908
Qy        928 GGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACT 987
Db        909 GGCGCATCCTGCTGTCGGGCATCGGCATCCACTTCTTCCAGCAGGCGTCGGGCATCGACT 968
Qy        988 CCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACT 1047
Db        969 CCGTCGTGCTCTACAGCCCGCGCGTGTTCAAGAGCGCCGGCATCACCGACGACAAACACC 1028
Qy       1048 TCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTT 1107
Db       1029 TCCTCGGCACCACCTGCGCCGTCGGTGTCACCAAGACGCTCTTCATCCTCGTGGCGACCT 1088
Qy       1108 TCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCC 1167
Db       1089 TCTTCCTCGACCGCGTCGGGCGGCGGCCGCTGCTGCTGAGCAGCACGGGCGGGATGATCC 1148
Qy       1168 TCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGA 1227
Db       1149 TCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGA 1208
Qy       1228 TACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCA 1287
Db       1209 TACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCA 1268
Qy       1288 TCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCG 1347
Db       1269 TCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCG 1328
Qy       1348 CGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGA 1407
Db       1329 CGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGA 1388
Qy       1408 CCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCA 1467
Db       1389 CCTTCCTGTCGCTGTCCAGGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCA 1448
Qy       1468 TCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGC 1527
Db       1449 TCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGC 1508
Qy       1528 TGGAGGAGATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAG 1587
Db       1509 TGGAGGAGATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAG 1568
Qy       1588 CCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCACTAACTGATCAAACTAACCGCAAAATC 1647
Db       1569 CCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCACTAACTGATCAAACTAACCGCAAAATC 1628
Qy       1648 ACCAAATCCTAAGGGTTTTCTTGCAAAAACGTGTGCTGTACTGGCTAGCTAGCAAGTAGT 1707
Db       1629 ACCAAATCCTAAGGGTTTTCTTGCAAAAACGTGTGCTGTACTGGCTAGCTAGCAAGTAGT 1688
Qy       1708 AGCAGCAACGTGGGAAGATTCGCTGATCCGGCGTTGCTGGAGAGCGACGGCCGGCGACGA 1767
Db       1689 AGCAGCAACGTGGGAAGATTCGCTGATCCGGCGTTGCTGGAGAGCGACGGCCGGCGACGA 1748
Qy       1768 CAAAGCTGAGCTCCAGCTCGAGACTTCTTAAAATCATCTTCAAGTACATGGATTTTATTT 1827
Db       1749 CAAAGCTGAGCTCCAGCTCGAGACTTCTTAAAATCATCTTCAAGTACATGGATTTTATTT 1808
Qy       1828 TGCTCTTTGCTTTGTCCGTAAAAGTTGTACTATGCGATGAAGAATACCAGTATGTAGCAA 1887
Db       1809 TGCTCTTTGCTTTGTCCGTAAAAGTTGTACTATGCGATGAAGAATACCAGTATGTAGCAA 1868
Qy       1888 GGCTGAGGTTGTGTGTAGCTACTAGAAGTGTCAGTCACGTTGTTCTTGTAAGAAATGTTT 1947
Db       1869 GGCTGAGGTTGTGTGTAGCTACTAGAAGTGTCAGTCACGTTGTTCTTGTAAGAAATGTTT 1928
Qy       1948 AACTGTTAATTAAGCAGTATTGTTGCAGTAATC 1980
Db       1929 AACTGTTAATTAAGCAGTATTGTTGCAGTAATC 1961
```

RESULT 3
AK059423

LOCUS	AK059423 1962 bp mRNA linear PLN 24-JUL-2003
DEFINITION	Oryza sativa (japonica cultivar-group) cDNA clone:001-027-E03, full insert sequence.
ACCESSION	AK059423
VERSION	AK059423.1 GI:32969441
KEYWORDS	FLI_CDNA; oligo-capping.
SOURCE	Oryza sativa (japonica cultivar-group)
ORGANISM	Oryza sativa (japonica cultivar-group) Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE	1
AUTHORS	The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team:, Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group:, Otomo,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y., Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M., Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J., Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J., Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN:, Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y., Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Yoshino,M. and Hayashizaki,Y.
TITLE	Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice
JOURNAL	Science 301 (5631), 376-379 (2003)
MEDLINE	22752273
PUBMED	12869764
REFERENCE	2 (bases 1 to 1962)
AUTHORS	Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K., Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K., Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M., Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M., Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A., Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M., Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K., Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H., Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K., Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S., Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y.,

Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W., Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and Yoshimura,A.

TITLE Direct Submission

JOURNAL Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)

COMMENT This clone is one of the 28K full-length cDNA clones from japonica rice.
URL : http://cdna01.dna.affrc.go.jp/cDNA/
NIAS Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and Yamamoto,M.
FAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y., Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M., Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J., Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S., Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S., Yoshimura,A., Matsubara,K. and Murakami,K.
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J., Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N., Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K., Yasunishi,A. and Hayashizaki,Y.

FEATURES Location/Qualifiers
source 1. .1962
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="001-027-E03"

ORIGIN

Query Match 93.2%; Score 1880.6; DB 8; Length 1962;
Best Local Similarity 97.6%; Pred. No. 1.3e-246;
Matches 1907; Conservative 1; Mismatches 45; Indels 0; Gaps 0;

```
Qy         28 GCTTACACTCGACCGCCACTACTGTACACGGCCCAGAGCGAGCCTCCTCCTCCTCTGCAC 87
              ||  | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         10 GCGAGCTCTCGACCGCCACTACTGTACACGGCCCAGAGCGAGCCTCCTCCTCCTCTGCAC 69

Qy         88 CACCGGAGATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACG 147
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         70 CACCGGAGATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACG 129

Qy        148 TCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACG 207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 TCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACG 189

Qy        208 ATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACG 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 ATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACG 249

Qy        268 GGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGG 327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 GGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGG 309

Qy        328 CGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATAT 387
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        310 CGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATAT 369

Qy        388 TCTTCGCGGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCC 447
              ||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        370 TCTTCGCGGGGGCGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCC 429

Qy        448 GCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCG 507
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        430 GCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCG 489

Qy        508 AGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACT 567
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        490 AGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACT 549

Qy        568 TCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCG 627
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        550 TCGGCATCCTGCTCGGGTACGTCTCGAACTATGCCTTCTCCCGCTTGCCGCTGAACCTCG 609

Qy        628 GGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGC 687
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        610 GGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGGGC 669

Qy        688 TCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGG 747
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        670 TCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGG 729

Qy        748 TGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCA 807
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        730 TGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCA 789

Qy        808 AGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAG 867
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        790 AGGCCGCCGCCGGCATCCCCGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAG 849

Qy        868 GGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGC 927
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        850 GGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGC 909

Qy        928 GGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACT 987
              |||||||||||||||| || ||||||| |||||||||||||| |||| ||||||  |||
Db        910 GGCGCATCCTGCTGTCGGGCATCGGCACCCACTTCTTCCAGCAGGCGTCGGGCATCGACT 969

Qy        988 CCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACT 1047
              ||||||| |||||||||| | ||||||||||||||| |||| | ||  ||||||||||
Db        970 CCGTCGTGCTCTACAGCCCGCGCGTGTTCAAGAGCGCCGGCATCACCGACGACAAACACC 1029

Qy       1048 TCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTT 1107
              || | ||||||||  ||  |  ||||||||||||||| ||| ||||||  |  |||||| |
Db       1030 TCCTCGGCACCACCTGCGCCGTCGGTGTCACCAAGACGCTCTTCATCCTCGTGGCGACCT 1089

Qy       1108 TCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCC 1167
              |||||  |||||  ||||||||||||||||||||  |||||  |||||||||||||| ||||
Db       1090 TCTTCCTCGACCGCGTCGGGCGGCGGCCGCTGCTGCTGAGCAGCACGGGCGGGATGATCC 1149

Qy       1168 TCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGA 1227
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1150 TCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGA 1209

Qy       1228 TACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCA 1287
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1210 TACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCA 1269

Qy       1288 TCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCG 1347
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1270 TCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGGGATCTTCCCGCTCCAGGTGCGCG 1329

Qy       1348 CGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGA 1407
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1330 CGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGA 1389

Qy       1408 CCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCA 1467
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1390 CCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCA 1449
```

```
Qy       1468 TCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGC 1527
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1450 TCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGC 1509

Qy       1528 TGGAGGAGATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAG 1587
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1510 TGGAGGAGATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAG 1569

Qy       1588 CCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCACTAACTGATCAAACTAACCGCAAAATC 1647
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1570 CCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCACTAACTGATCAAACTAACCGCAAAATC 1629

Qy       1648 ACCAAATCCTAAGGGTTTTCTTGCAAAAACGTGTGCTGTACTGGCTAGCTAGCAAGTAGT 1707
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1630 ACCAAATCCTAAGGGTTTTCTTGCAAAAACGTGTGCTGTACTGGCTAGCTAGCAAGTAAT 1689

Qy       1708 AGCAGCAACGTGGGAAGATTCGCTGATCCGGCGTTGCTGGAGAGCGACGGCCGGCGACGA 1767
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1690 AGCAGCAACGTGGGAAGATTCGCTGATCCGGCGTTGCTGGAGAGCGACGGCCGGCGACGA 1749

Qy       1768 CAAAGCTGAGCTCCAGCTCGAGACTTCTTAAAATCATCTTCAAGTACATGGATTTTATTT 1827
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1750 CAAAGCTGAGCTCCAGCTCGAGACTTCTTAAAATCATCTTCAAGTACATGGATTTTATTT 1809

Qy       1828 TGCTCTTTGCTTTGTCCGTAAAAGTTGTACTATGCGATGAAGAATACCAGTATGTAGCAA 1887
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1810 TGCTCTTTGCTTTGTCCGTAAAAGTTGTACTATGCGATGAAGAATACCAGTATGTAGCAA 1869

Qy       1888 GGCTGAGGTTGTGTGTAGCTACTAGAAGTGTCAGTCACGTTGTTCTTGTAAGAAATGTTT 1947
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1870 GGCTGAGGTTGTGTGTAGCTACTAGAAGTGTCAGTCACGTTGTTCTTGTAAGAAATGTTT 1929

Qy       1948 AACTGTTAATTAAGCAGTATTGTTGCAGTAATC 1980
             |||||||||||||||||||||||||||||||
Db       1930 AACTGTTAATTAAGCAGTATTGTTGCAGTAATC 1962
```

RESULT 4
AP004185
LOCUS AP004185 25270 bp DNA linear PLN 02-SEP-2003
DEFINITION Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 7, BAC clone:OJ1301_C12.
ACCESSION AP004185
VERSION AP004185.3 GI:32261236
KEYWORDS .
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1
AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.
TITLE Oryza sativa nipponbare(GA3) genomic DNA, chromosome 7, BAC clone:OJ1301_C12
JOURNAL Published Only in Database (2001)
REFERENCE 2 (bases 1 to 25270)
AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.
TITLE Direct Submission
JOURNAL Submitted (20-SEP-2001) Takuji Sasaki, National Institute of Agrobiological Sciences, Rice Genome Research Program; Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan (E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/, Tel:81-298-38-7441, Fax:81-298-38-7468)
COMMENT On Jun 25, 2003 this sequence version replaced gi:21280394.
Genes were predicted from the integrated results of the following: GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), FGENESH (http://www.softberry.com/), GeneMark.hmm (http://opal.biology.gatech.edu/GeneMark/), GlimmerM (http://www.tigr.org/tdb/glimmerm/glmr_form.html), RiceHMM (http://rgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor (http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4 (http://globin.cse.psu.edu/html/docs/sim4.html), gap2 (http://www.tigr.org/software/glimmerm/), BLASTN and BLASTX. The genomic sequence was searched against NCBI NonRedundant Protein database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at RGP or DDBJ. Protein homologies of the coding regions were searched against NCBI NonRedundant Protein database with BLASTP. ESTs represent the identified cDNA sequences using BLASTN with the corresponding DDBJ accession no. and RGP clone ID. Full-length cDNAs represent the identified cDNA sequences using BLASTN with the corresponding DDBJ accession no.
A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with full-length cDNA or EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted by two or more gene prediction programs is classified as a 'hypothetical' protein according to IRGSP standard. A gene predicted by a single gene prediction program is also classified as a probable 'hypothetical' protein and is included as a miscellaneous feature of the sequence.
The orientation of the sequence is from -21M13 to M13rev of the BAC clone. This is a partial sequence of OJ1301_C12 clone.This sequence of OJ1301_C12 clone has an overlap with P0453G03 clone(DDBJ: AP004276)at 5' end and an overlap with OJ1127_E01(DDBJ: AP003747) at 3' end.The sequence was generated by combining Monsanto and RGP-Japan sequencing data. Detailed information on overlap and assembly quality together with annotation of this entry is available at http://rgp.dna.affrc.go.jp/GenomeSeq.html.

```
FEATURES             Location/Qualifiers
     source          1. .25270
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="genomic DNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /chromosome="7"
                     /clone="OJ1301_C12"
     gene            779. .1103
                     /gene="OJ1301_C12.1"
     mRNA            join(<779. .940,1011. .>1103)
                     /gene="OJ1301_C12.1"
                     /note="start and end point are not identified"
     CDS             join(779. .940,1011. .1103)
                     /gene="OJ1301_C12.1"
                     /note="predicted by GENSCAN etc."
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="BAC83309.1"
                     /db_xref="GI:34393629"
                     /translation="MYRERAQRRGRSSATVFLASSCRQLGKGKKVFARDGNPAGCGYP
                     CRPASARICTLIRLSADWTTAACDDGSSAAVDFHGSFRFS"
     gene            complement(join(1344. .1480,2498. .2686,3144. .3198))
                     /gene="OJ1301_C12.2"
     misc_feature    complement(join(1344. .1480,2498. .2686,3144. .3198))
                     /gene="OJ1301_C12.2"
                     /note="hypothetical ORF
                     predicted by GENSCAN
                     this category is not included in IRGSP standard"
     gene            5120. .7180
                     /gene="OJ1301_C12.3"
     mRNA            join(5120. .5308,5408. .7180)
                     /gene="OJ1301_C12.3"
                     /note="supported by full-length cDNA(s): AK059423"
     mRNA            join(5121. .5308,5408. .7180)
                     /gene="OJ1301_C12.3"
                     /note="supported by full-length cDNA(s): AK102676"
     CDS             join(5197. .5308,5408. .6828)
                     /gene="OJ1301_C12.3"
                     /note="contains EST(s):
                     C20132(E11563),AU163952(E11563),AU097469(S5067),
                     AU033198(S5067)
                     contains full-length cDNA(s): AK059423,AK102676"
                     /codon_start=1
```

```
                /product="putative sorbitol transporter"
                /protein_id="BAC83310.1"
                /db_xref="GI:34393630"
                /translation="MASAALPEAVAPKKKGNVRFAFACAILASMTSILLGYDIGVMSG
                ASLYIKKDFNISDGKVEVLMGILNLYSLIGSFAAGRTSDWIGRRYTIVFAAVIFFAGA
                FLMGFAVNYAMLMFGRFVAGIGVGYALMIAPVYTAEVSPASARGFLTSFPEVFINFGI
                LLGYVSNYAFSRLPLNLGWRIMLGIGAAPSVLLALMVLGMPESPRWLVMKGRLADAKV
                VLEKTSDTAEEAAERLADIKAAAGIPEELDGDVVTVPKRGSGNEKRVWKELILSPTPA
                MRRILLSGIGIHFFQQASGIDSVVLYSPRVFKSAGITDDKHLLGTTCAVGVTKTLFIL
                VATFFLDRVGRRPLLLSSTGGMILSLIGLGAGLTVVGQHPDAKIPWAIGLSIASTLAY
                VAFFSIGLGPITWVYSSEIFPLQVRALGCSLGVAANRVTSGVISMTFLSLSKAITIGG
                SFFLYSGIAALAWVFFYTYLPETRGRTLEEMSKLFGDTAAASESDEPAKEKKKVEMAA
                TN"
gene            complement(7647. .9324)
                /gene="OJ1301_C12.4"
mRNA            complement(join(<7647. .9067,9210. .>9324))
                /gene="OJ1301_C12.4"
                /note="start and end point are not identified"
CDS             complement(join(7647. .9067,9210. .9324))
                /gene="OJ1301_C12.4"
                /codon_start=1
                /product="putative sorbitol transporter"
                /protein_id="BAC83311.1"
                /db_xref="GI:34393631"
                /translation="MPESTKLTAMAVDPKKKNASYAFTCAILASMASIILGYDIGVMS
                GASLYIKKDLKITDVQVEILMGILNIYSLVGSFAAGRTADWIGRRFTVVFAAAFFFAR
                ALLMGFSGDYATLMVGRFVAGVGVGYAIMIAPVYTAEISPASSRGFLTSFPEVSINLG
                ILLGYVSNYAFARLPLSLGWRVMLGVGAAPSVLLALMVLGMPESPRWLVMKGRLADAK
                AVLEKIADTPEEASERLADIKAAAGIPDDLDGDVVTVSKKRGGEEGQVWRELVVSPTP
                AMRRIVLAAVGLHFFQQASGVDSVVLYSPRVFQSAGITGDDQLLGTTCAVGFAKTVFI
                LVAAFLLDRAGRRPLLLTSTGGMVFSLVGLATGLTVVGGSPDAQVPSWAVGLCVASIL
                AYVAFFSVGLGPMSGVYTSEIFPLRARALGFAVAVACNRVTSGVISMTFLSLSSAITI
                GGSFFLYAAISSLAWVFFFTRLPETRGQTLEEIGKVFGMDDTAMEAEDSAAYRERLLA
                TSP"
gene            11758. .12731
                /gene="OJ1301_C12.5"
mRNA            join(<11758. .11894,11990. .12071,12243. .>12731)
                /gene="OJ1301_C12.5"
                /note="start and end point are not identified"
CDS             join(11758. .11894,11990. .12071,12243. .12731)
                /gene="OJ1301_C12.5"
                /note="predicted by FGENESH etc."
                /codon_start=1
                /product="hypothetical protein"
                /protein_id="BAC83312.1"
                /db_xref="GI:34393632"
                /translation="MAKRARTLTGDHLRSSSSSNQAPSWGAKRARSGSALRDRSLPDS
                EWRDWSSIGNGPAWLISERVLAGDDVADYADLPELRGHRVFGPTAEGLLVLLDEATFA
                VRLLNPLTRQLTDLPPATTVLDPTSLMISPVHYLLHVLTVHGAGLADGSATVVVYFEN
                ARKLAAAKPGDERWTVVEPDASFASVSSFAGRFYCATRHAIRPVIVEVRTTDHRLPPQ
                LVVVAKLAGVPSFPRNC"
gene            join(12995. .17652,18451. .19187)
                /gene="OJ1301_C12.6"
repeat_unit     12995. .13731
                /gene="OJ1301_C12.6"
                /note="5' Terminal repeat"
                /rpt_type=terminal
misc_feature    13837. .17652
                /gene="OJ1301_C12.6"
                /note="probably inactive due to including stop codon(s) in
                CDS
                pseudogene,Fourf gag/pol protein"
                /pseudo
repeat_unit     18451. .19187
                /gene="OJ1301_C12.6"
                /note="3' Terminal repeat"
                /rpt_type=terminal
gene            complement(19137. .20909)
                /gene="OJ1301_C12.7"
mRNA            complement(19137. .20909)
                /gene="OJ1301_C12.7"
                /note="supported by full-length cDNA(s): AK110773"
gene            19581. .21159
                /gene="OJ1301_C12.8"
mRNA            19581. .21159
                /gene="OJ1301_C12.8"
                /note="supported by full-length cDNA(s): AK069022"
CDS             complement(19794. .20309)
                /gene="OJ1301_C12.7"
                /note="contains full-length cDNA(s): AK110773"
                /codon_start=1
                /product="unknown protein"
                /protein_id="BAC83313.1"
                /db_xref="GI:34393633"
                /translation="MPSSAPASQRSHRAHAGLTASKPPPTARSFRWPHLPPCRPRQPP
                QPPSAPPRQPTLPPGVSPLTARASSRAPCRPPPPPAVPATCCLLPRLPPPASCYHLRR
                CELRRSETCLAMTSASSASRRTAPRVPPPPESPAPDPATGRPDQPPRGPNPPPAWPDS
                PCGPLRTPSSA"
CDS             20013. .20480
                /gene="OJ1301_C12.8"
                /note="contains EST(s):
                D40430(S2406),D15710(C1130A),AU086036(S2406)
                contains full-length cDNA(s): AK069022"
                /codon_start=1
```

Query Match 88.4%; Score 1783; DB 8; Length 25270;
Best Local Similarity 93.1%; Pred. No. 2.3e-233;
Matches 1916; Conservative 1; Mismatches 41; Indels 99; Gaps 1;

```
Qy      28 GCTTACACTCGACCGCCACTACTGTACACGGCCCAGAGCGAGCCTCCTCCTCCTCTGCAC 87
           ||  | ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5129 GCGAGCTCTCGACCGCCACTACTGTACACGGCCCAGAGCGAGCCTCCTCCTCCTCTGCAC 5188

Qy      88 CACCGGAGATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACG 147
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5189 CACCGGAGATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACG 5248

Qy     148 TCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTAC- 206
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5249 TCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACG 5308

Qy     207 ------------------------------------------------------------ 206

Db    5309 GTAACTAATCAAAGCGCATTCACTTCCCGCCCAGCACAGCACGAGCTGCTGCTTCTCCCC 5368

Qy     207 ----------------------------------------GATATCGGGGTGATGAGCGGGG 228
                                                     |||||||||||||||||||||
Db    5369 AGCATGGGCGCTCATGGCGATGTGCATGTGACCTTGCAGATATCGGGGTGATGAGCGGGG 5428

Qy     229 CGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGG 288
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5429 CGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGG 5488

Qy     289 GCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGA 348
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5489 GCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGA 5548

Qy     349 TCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGSGTTCCTCA 408
           |||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db    5549 TCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGCGTTCCTCA 5608

Qy     409 TGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCG 468
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5609 TGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCG 5668

Qy     469 TGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGC 528
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5669 TGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGC 5728

Qy     529 GTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACG 588
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5729 GTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACG 5788

Qy     589 TCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCA 648
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5789 TCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCA 5848
```

```
Qy       649 TCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGC 708
Db      5849 TCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGC 5908

Qy       709 GGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCG 768
Db      5909 GGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCG 5968

Qy       769 ACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTG 828
Db      5969 ACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCCG 6028

Qy       829 AGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGG 888
Db      6029 AGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGG 6088

Qy       889 TGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGA 948
Db      6089 TGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCGGGCA 6148

Qy       949 TCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTC 1008
Db      6149 TCGGCATCCACTTCTTCCAGCAGGCGTCGGGCATCGACTCCGTCGTGCTCTACAGCCCGC 6208

Qy      1009 TCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCGT 1068
Db      6209 GCGTGTTCAAGAGCGCCGGCATCACCGACGACAAACACCTCCTCGGCACCACCTGCGCCG 6268

Qy      1069 TCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGC 1128
Db      6269 TCGGTGTCACCAAGACGCTCTTCATCCTCGTGGCGACCTTCTTCCTCGACCGCGTCGGGC 6328

Qy      1129 GGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCG 1188
Db      6329 GGCGGCCGCTGCTGCTGAGCAGCACGGGCGGGATGATCCTCTCCCTCATCGGCCTCGGCG 6388

Qy      1189 CCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAA 1248
Db      6389 CCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAA 6448

Qy      1249 GCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGT 1308
Db      6449 GCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGT 6508

Qy      1309 GGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCG 1368
Db      6509 GGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCG 6568

Qy      1369 TCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGG 1428
Db      6569 TCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGG 6628

Qy      1429 CCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGT 1488
Db      6629 CCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGT 6688

Qy      1489 TCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGT 1548
Db      6689 TCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGT 6748

Qy      1549 TCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAGCCAAGGAGAAGAAGAAGGTGG 1608
Db      6749 TCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAGCCAAGGAGAAGAAGAAGGTGG 6808

Qy      1609 AAATGGCCGCCACTAACTGATCAAACTAACCGCAAAATCACCAAATCCTAAGGGTTTTCT 1668
Db      6809 AAATGGCCGCCACTAACTGATCAAACTAACCGCAAAATCACCAAATCCTAAGGGTTTTCT 6868

Qy      1669 TGCAAAAACGTGTGCTGTACTGGCTAGCTAGCAAGTAGTAGCAGCAACGTGGGAAGATTC 1728
Db      6869 TGCAAAAACGTGTGCTGTACTGGCTAGCTAGCAAGTAGTAGCAGCAACGTGGGAAGATTC 6928

Qy      1729 GCTGATCCGGCGTTGCTGGAGAGCGACGGCCGGCGACGACAAAGCTGAGCTCCAGCTCGA 1788
Db      6929 GCTGATCCGGCGTTGCTGGAGAGCGACGGCCGGCGACGACAAAGCTGAGCTCCAGCTCGA 6988

Qy      1789 GACTTCTTAAAATCATCTTCAAGTACATGGATTTTATTTTGCTCTTTGCTTTGTCCGTAA 1848
Db      6989 GACTTCTTAAAATCATCTTCAAGTACATGGATTTTATTTTGCTCTTTGCTTTGTCCGTAA 7048

Qy      1849 AAGTTGTACTATGCGATGAAGAATACCAGTATGTAGCAAGGCTGAGGTTGTGTGTAGCTA 1908
Db      7049 AAGTTGTACTATGCGATGAAGAATACCAGTATGTAGCAAGGCTGAGGTTGTGTGTAGCTA 7108

Qy      1909 CTAGAAGTGTCAGTCACGTTGTTCTTGTAAGAAATGTTTAACTGTTAATTAAGCAGTATT 1968
Db      7109 CTAGAAGTGTCAGTCACGTTGTTCTTGTAAGAAATGTTTAACTGTTAATTAAGCAGTATT 7168

Qy      1969 GTTGCAGTAATCAAAAA 1985
Db      7169 GTTGCAGTAATCAACAA 7185
```

RESULT 5
AF416867
LOCUS AF416867 1860 bp mRNA linear PLN 15-OCT-2001
DEFINITION Oryza sativa putative sugar transporter mRNA, complete cds.
ACCESSION AF416867
VERSION AF416867.1 GI:16118826
KEYWORDS .
SOURCE Oryza sativa
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 1860)
AUTHORS Wu,P., Chen,Q., Huang,G. and Yi,K.
TITLE Molecular cloning of putative sugar transporter in rice
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1860)
AUTHORS Wu,P., Chen,Q., Huang,G. and Yi,K.
TITLE Direct Submission
JOURNAL Submitted (06-SEP-2001) College of Life Science, Plant Molecular Physiology and Functional Genomics Laboratory, 268 Kai Xuan Rd, Hangzhou, Zhejiang 310029, P.R. China
FEATURES Location/Qualifiers
source 1. .1860
/organism="Oryza sativa"
/mol_type="mRNA"
/db_xref="taxon:4530"
CDS 3. .1736
/codon_start=1
/product="putative sugar transporter"
/protein_id="AAL14615.1"
/db_xref="GI:16118827"
/translation="MSATTSNSRRLCNGRGQVIPGPTLPEAVKMGPHAIRIVVGGLGE
GELSTATTVHGPERASSSSAPPEMASAALPEAVAPKKKGNVRFAFACAILASMTSILL
GYDIGVMSGASLYIKKDFNISDGKVEVLMGILNLYSLIGSFAAGRTSDWIGRRYTIVF
AAVIFFAGAFLMGFAVNYAMLMFGRFVAGIGVGYALMIAPVYTAEVSPASARGFLTSF
PEVFINFGILLGYVSNYAFSRLPLNLGWRIMLGIGAAPSVLLALMVLGMPESPRWLVM
KGRLADAKVVLEKTSDTAEEAAERLADIKAAAGIPEELDGDVVTVPKRGSGNEKRVWK
ELILSPTPAMRRILLSGIGIHFFQQASGIDSVVLYSPRVFKSAGITDDKHLLGTTCAV
GVTKTLFILVATFFLDRVGRRPLLLSSTGGMILSLIGLGAGLTVVGQHPDAKIPWAIG
LSIASTLAYVAFFSIGLGPITWVYSSEIFPLQVRALGCSLGVAANRVTSGVISMTFLS
LSKAITIGGSFFLYSGIAALAWVFFYTYLPETRGRTLEEMSKLFGDTAAASESDEPAK
EKKKVEMAATN"
ORIGIN

Query Match 82.3%; Score 1660.6; DB 8; Length 1860;
Best Local Similarity 97.6%; Pred. No. 1.2e-216;
Matches 1684; Conservative 1; Mismatches 40; Indels 0; Gaps 0;

```
Qy        28 GCTTACACTCGACCGCCACTACTGTACACGGCCCAGAGCGAGCCTCCTCCTCCTCTGCAC 87
Db       136 GCGAGCTCTCGACCGCCACTACTGTACACGGCCCAGAGCGAGCCTCCTCCTCCTCTGCAC 195
```

```
Qy        88 CACCGGAGATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACG 147
Db       196 CACCGGAGATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACG 255

Qy       148 TCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACG 207
Db       256 TCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACG 315

Qy       208 ATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACG 267
Db       316 ATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACG 375

Qy       268 GGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGG 327
Db       376 GGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGG 435

Qy       328 CGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATAT 387
Db       436 CGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATAT 495

Qy       388 TCTTCGCGGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCC 447
Db       496 TCTTCGCGGGGGCGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCC 555

Qy       448 GCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCG 507
Db       556 GCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCG 615

Qy       508 AGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACT 567
Db       616 AGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACT 675

Qy       568 TCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCG 627
Db       676 TCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCG 735

Qy       628 GGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGC 687
Db       736 GGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGC 795

Qy       688 TCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGG 747
Db       796 TCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGG 855

Qy       748 TGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCA 807
Db       856 TGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCA 915

Qy       808 AGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAG 867
Db       916 AGGCCGCCGCCGGCATCCCCGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAG 975

Qy       868 GGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGC 927
Db       976 GGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGC 1035

Qy       928 GGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACT 987
Db      1036 GGCGCATCCTGCTGTCGGGCATCGGCATCCACTTCTTCCAGCAGGCGTCGGGCATCGACT 1095

Qy       988 CCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACT 1047
Db      1096 CCGTCGTGCTCTACAGCCCGCGCGTGTTCAAGAGCGCCGGCATCACCGACGACAAACACC 1155

Qy       1048 TCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTT 1107
Db       1156 TCCTCGGCACCACCTGCGCCGTCGGTGTCACCAAGACGCTCTTCATCCTCGTGGCGACCT 1215

Qy       1108 TCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCC 1167
Db       1216 TCTTCCTCGACCGCGTCGGGCGGCGGCCGCTGCTGCTGAGCAGCACGGGCGGGATGATCC 1275

Qy       1168 TCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGA 1227
Db       1276 TCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGA 1335

Qy       1228 TACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCA 1287
Db       1336 TACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCA 1395

Qy       1288 TCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCG 1347
Db       1396 TCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCG 1455

Qy       1348 CGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGA 1407
Db       1456 CGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGA 1515

Qy       1408 CCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCA 1467
Db       1516 CCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCA 1575

Qy       1468 TCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGC 1527
Db       1576 TCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGC 1635

Qy       1528 TGGAGGAGATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAG 1587
Db       1636 TGGAGGAGATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAG 1695

Qy       1588 CCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCACTAACTGATCAAACTAACCGCAAAATC 1647
Db       1696 CCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCACTAACTGATCAAACTAACCGCAAAATC 1755

Qy       1648 ACCAAATCCTAAGGGTTTTCTTGCAAAAACGTGTGCTGTACTGGCTAGCTAGCAAGTAGT 1707
Db       1756 ACCAAATCCTAAGGGTTTTCTTGCAAAAACGTGTGCTGTACTGGCTAGCTAGCAAGTAGT 1815

Qy       1708 AGCAGCAACGTGGGAAGATTCGCTGATCCGGCGTTGCTGGAGAGC 1752
Db       1816 AGCAGCAACGTGGGAAGATTCGCTGATCCGGCGTTGCTGGAGAGC 1860
```

RESULT 6
BT009264
LOCUS BT009264 2109 bp mRNA linear PLN 20-JUN-2003
DEFINITION Triticum aestivum clone wlk8.pk0001.a11:fis, full insert mRNA sequence.
ACCESSION BT009264
VERSION BT009264.1 GI:32128815
KEYWORDS FLI_CDNA.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 2109)
AUTHORS Tingey,S.V., Wolters,P., Powell,W., Dolan,M., Miao,G.-H., Caraher,N.R., Hanafey,M.K. and Hainey,C.F.
TITLE Direct Submission
JOURNAL Submitted (20-JUN-2003) Crop Genetics, E. I. DuPont de Nemours and Company, 1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
FEATURES Location/Qualifiers
source 1. .2109
/organism="Triticum aestivum"
/mol_type="mRNA"
/db_xref="taxon:4565"
/clone="wlk8.pk0001.a11:fis"
ORIGIN

Query Match 55.5%; Score 1120.2; DB 8; Length 2109;
Best Local Similarity 80.1%; Pred. No. 4.3e-143;
Matches 1359; Conservative 1; Mismatches 319; Indels 18; Gaps 3;

```
Qy        84 GCACCACCGGAGATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGC 143
Db        87 GCGTCGTCGACGATGGACCGCGCCGCACTCCCGGCGGCCGTCGAGCCCAAGAAGAAGGGC 146

Qy       144 AACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGC 203
Db       147 AACGTGAGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGC 206

Qy       204 TACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGT 263
Db       207 TACGACATCGGCGTGATGAGCGGAGCGTCGCTGTACATCCAGAAGGATCTGAAGATCAAC 266

Qy       264 GACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTC 323
Db       267 GACACCCAGCTGGAGGTCCTCATGGGCATCCTCAACGTGTACTCGCTCATTGGCTCCTTC 326

Qy       324 GCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTC 383
Db       327 GCGGCGGGGCGGACGTCCGACTGGATCGGCCGGCGCTTCACCATCGTCTTCGCCGCCGTC 386

Qy       384 ATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTC 443
Db       387 ATCTTCTTCGCGGGCGCCCTCATCATGGGCTTCTCCGTCAACTACGCCATGCTCATGTTC 446

Qy       444 GGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACC 503
Db       447 GGCCGCTTCGTGGCCGGCATCGGCGTGGGGTACGCTCTCATGATCGCGCCCGTGTACACG 506

Qy       504 GCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATC 563
Db       507 GCCGAGGTGTCCCCCGCGTCTGCCCGTGGCTTTCTCACCTCCTTCCCGGAGGTGTTCATC 566

Qy       564 AACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAAC 623
Db       567 AACTTCGGCATCCTCCTCGGATATGTCTCCAACTTCGCCTTCGCCCGCCTCTCCCTCCGC 626

Qy       624 CTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATG 683
Db       627 CTCGGCTGGCGCATTATGCTCGGCATAGGCGCGGTGCCGTCCGTCCTGCTCGCCTTCATG 686

Qy       684 GTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCC 743
Db       687 GTGCTCGGCATGCCCGAGTCTCCCCGGTGGCTCGTCATGAAGGGCCGTCTCGCGGACGCC 746

Qy       744 AAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGAC 803
Db       747 AAGGTTGTGCTTGCCAAGACTTCCGACACGCCGGAGGAGGCCGCCGAGCGCCTCGCCGAC 806

Qy       804 ATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCC--- 860
Db       807 ATTAAGACTGCCGCCGGCATCCCTCTGGACCTCGACGGCGACGTGGTCGCCGTGCCCAAA 866

Qy       861 AAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCG 920
Db       867 AACAAAGGAAGCAGCGAGGAGAAGCGCGTGTGGAAGGACCTCATCCTGTCACCGACCAAA 926

Qy       921 GCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGC 980
Db       927 GCCATGCGCCACATCCTCATCGCGGGAATCGGCATCCACTTCTTCCAGCAGTCTTCGGGC 986

Qy       981 ATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGAC 1040
Db       987 ATCGACGCCGTCGTGCTCTACAGCCCGCTAGTTTTCAAGAGCGCCGGCATCACGGGCGAC 1046

Qy      1041 AAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTG 1100
Db      1047 AACCGTCTCCTCGGCACCACCGTGGCGGTCGGGGCCACCAAGACGGTCTTCATCCTGGTG 1106

Qy      1101 GCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGG 1160
Db      1107 GCCACCTTCCTCCTCGACCGCATCGGCCGGCGGCCGCTGCTGCTGAGCAGCACGGGCGGC 1166

Qy      1161 ATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGAC 1220

Db      1167 ATGATCGTCTCCCTAGTGGGCCTCGCGACGGGGCTCACCGTCATCAGCCGCCACCCGGAC 1226

Qy      1221 GCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTC 1280
Db      1227 GAGAAGATCACCTGGGCCATCGTCCTGTGCATCTTCTGCATCATGGCCTACGTGGCCTTC 1286

Qy      1281 TTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAG 1340
Db      1287 TTCTCCATCGGCCTCGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTGCAC 1346

Qy      1341 GTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATC 1400
Db      1347 GTGCGCGCGCTGGGCTGCTCCCTGGGCGTGGCCGTCAACCGCCTGACCAGCGGCGTGATC 1406

Qy      1401 TCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTAC 1460
Db      1407 TCCATGACCTTCATTTCGCTGTCCAAGGCCATGACCATCGGCGGCGCCTTCTTCCTCTTC 1466

Qy      1461 TCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGC 1520
Db      1467 GCCGGCATCGCCTCATTCGCATGGGTGTTCTTCTTCGCCTACCTGCCGGAGACCCGCGGC 1526

Qy      1521 CGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGAC 1580
Db      1527 CGCACGCTGGAGGACATGAGCTCGCTGTTCGGCAACACGGCCACGCACAAGCAGGGCGCC 1586

Qy      1581 GAGCCAGCCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCACTAACTGATCAAACTAACCG 1640
Db      1587 GCGGAAGCCGACGACGACGCCGGGGAGAAGAAGGTGGAAATGGCCGCCACCAACTGACCG 1646

Qy      1641 CAAAATCACCAAATCCTAAGGGTTTTCTTGCAAAAACGTGTGCTGTACTGGCTAGCTAGC 1700
Db      1647 CAAGTTGGCAGATCGCGATGCGAAGACTTGCGCTGTATCCGTCTCGGCTAGCTAGCTGCC 1706

Qy      1701 A-----------AGTAGTAGCAGCAACGTGGGAAGATTCGCTGATCCGGC----GTTGCT 1745
Db      1707 ACAAGGCCACATAGATGACGAAGTAGCGTGGGAAGATTCGCTGATCCGGCCGGAGCTGCC 1766

Qy      1746 GGAGAGCGACGGCCGGC 1762
Db      1767 GGAGGGCGACGGCAAGC 1783
```

RESULT 7
AR208577

```
LOCUS       AR208577                2089 bp    DNA     linear   PAT 20-JUN-2002
DEFINITION  Sequence 25 from patent US 6383776.
ACCESSION   AR208577
VERSION     AR208577.1  GI:21509767
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2089)
  AUTHORS   Allen,S.M., Hitz,W.D., Kinney,A.J. and Tingey,S.V.
  TITLE     Plant sugar transport proteins
  JOURNAL   Patent: US 6383776-A 25 07-MAY-2002;
FEATURES             Location/Qualifiers
     source          1. .2089
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN
```

Query Match 54.3%; Score 1094.6; DB 6; Length 2089;
Best Local Similarity 79.1%; Pred. No. 1.3e-139;
Matches 1343; Conservative 1; Mismatches 335; Indels 18; Gaps 3;

```
Qy         84 GCACCACCGGAGATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGC 143
Db         65 GCGTCGTCGACGATGGACCGCGCCGCACTCCCGGCGGCCGTCGAGCCCAAGAAGAAGGGC 124

Qy        144 AACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGC 203
```

```
Db        125 AACGTGAGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGC 184
Qy        204 TACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGT 263
Db        185 TACGACATCGGCGTGATGAGCGGAGCGTCGCTGTACATCCAGAAGGATCTGAAGATCAAC 244
Qy        264 GACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTC 323
Db        245 GACACCCAGCTGGAGGTCCTCATGGGCATCCTCAACGTGTACTCGCTCATTGGCTCCTTC 304
Qy        324 GCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTC 383
Db        305 GCGGCGGGGCGGACGTCCGACTGGATCGGCCGGCGCTTCACCATCGTCTTCGCCGCCGTC 364
Qy        384 ATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTC 443
Db        365 ATCTTCTTCGCGGGCGCCCTCATCATGGGCTTCTCCGTCAACTACGCCATGCTCATGTTC 424
Qy        444 GGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACC 503
Db        425 GGGCGCTTCGTGGCCGGCATCGGCGTGGGGTACGCTCTCATGATCGCGCCCGTGAACACG 484
Qy        504 GCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATC 563
Db        485 GGCGAGGTGTCCCCCGCGTCTGCCCGTGGGGTTCTCACATCCTTCCCGGAGGTGTTCATC 544
Qy        564 AACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAAC 623
Db        545 AACTTCGGCATCCTCCTCGGATATGTCTCCAACTTCGCCTTCGCCCGCCTCTCCCTCCGC 604
Qy        624 CTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATG 683
Db        605 CTCGGCTGGCGCATTATGCTCGGCATAGGCGCGGTGCCGTCCGTCCTGCTCGCGTTCATG 664
Qy        684 GTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCC 743
Db        665 GTGCTCGGCATGCCCGAGTCTCCCCGGTGGCTCGTCATGAAGGGCCGTCTCGCGGACGCC 724
Qy        744 AAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGAC 803
Db        725 AAGGTTGTGCTTGCCAAGACGTCCGACACGCCGGAAGAGGCCGCCGAGCGCATCGCCGAC 784
Qy        804 ATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCC--- 860
Db        785 ATTAAGACTGCCGCCGGCATCCCTCTGGGCCTCGACGGCGACGTGGTCCCCGTGCCCAAA 844
Qy        861 AAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCG 920
Db        845 AACAAAGGAAGCAGCGAGGAGAAGCGCGTTTTGAAGGACCTCATCCTGTCACCGACCATA 904
Qy        921 GCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGC 980
Db        905 GCCATGCGCCACATCCTCATCGCGGGAATCGGCATCCACTTCTTCCAGCAGTCTTCGGGC 964
Qy        981 ATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGAC 1040
Db        965 ATCGACGCCGTCGTGCTCTACAGCCCGCTAGTTTTCAAGAGCGCCGGCATCACGGGCGAC 1024
Qy       1041 AAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTG 1100
Db       1025 AGCCGTCTCCGCGGCACCACCGTGGCGGTCGGGGCCACCAATACGGTCTTCATCCTGGTG 1084
Qy       1101 GCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGG 1160
Db       1085 GCCACCTTCCTCCTCGACCGCATCCGCCGGCGGCCGCTGGTGCTGACCAGCACGGGCGGC 1144
Qy       1161 ATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGAC 1220
Db       1145 ATGCTCGTCTCCTTAGTGGGCCTCGCGACGGGGCTCACCGTCATCAGCCGCCACCCGGAC 1204
Qy       1221 GCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTC 1280
```

```
Db        1205 GAGAAGATCACCTGGGCCATCGTCCTGTGCATCTTCTGCATCATGGCCTACGTGGCCTTC 1264
Qy        1281 TTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAG 1340
Db        1265 TTCTCCATCGGCCTCGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTGCAC 1324
Qy        1341 GTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATC 1400
Db        1325 GTGCGCGCGCTGGGCTGCTCCCTGGGCGTGGCCGTCAACCGCCTGACCAGCGGCGTGATC 1384
Qy        1401 TCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTAC 1460
Db        1385 TCCATGACCTTCATTTCGCTGTCCAAGGCCATGACCATCGGCGGCGCCTTCTTCCTCTTC 1444
Qy        1461 TCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGC 1520
Db        1445 GCCGGCATCGCCTCATTCGCATGGGTGTTCTTCTTCGCCTACCTGCCGGAGACCCGCGGC 1504
Qy        1521 CGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGAC 1580
Db        1505 CGCACGCTGGAGGACATGAGCTCGCTGTTCGGCAACACGGCCACGCACAAGCAGGGCGCC 1564
Qy        1581 GAGCCAGCCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCACTAACTGATCAAACTAACCG 1640
Db        1565 GCGGAAGCCGACGACGACGCCGGGGAGAAGAAGGTGGAAATGGCCGCCACCAACTGACCG 1624
Qy        1641 CAAAATCACCAAATCCTAAGGGTTTTCTTGCAAAAACGTGTGCTGTACTGGCTAGCTAGC 1700
Db        1625 CAAGTTGGCAGATCGCGATGCGAAGACTTGCGCTGTATCCGTCTCGGCTAGCTAGCTGCC 1684
Qy        1701 A-----------AGTAGTAGCAGCAACGTGGGAAGATTCGCTGATCCGGC----GTTGCT 1745
Db        1685 ACAAGGCCACATAGATGACGAAGTAGCGTGGGAAGATTCGCTGATCCGGCCGGAGCTGCC 1744
Qy        1746 GGAGAGCGACGGCCGGC 1762
Db        1745 GGAGGGCGACGGCAAGC 1761
```

RESULT 8
AK119464

```
LOCUS       AK119464                1620 bp    mRNA    linear   PLN 29-OCT-2003
DEFINITION  Oryza sativa (japonica cultivar-group) cDNA clone:001-133-F04, full
            insert sequence.
ACCESSION   AK119464
VERSION     AK119464.1  GI:37989087
KEYWORDS    FLI_CDNA; oligo capping.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1
  AUTHORS   The Rice Full-Length cDNA Consortium, National Institute of
            Agrobiological Sciences Rice Full-Length cDNA Project Team:,
            Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,
            Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,
            Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,
            Ohtsuki,K., Shishiki,T., Foundation of Advancement of International
            Science Genome Sequencing & Analysis Group:, Otomo,Y., Murakami,K.,
            Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,
            Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,
            Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J.,
            Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,
            Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN:,
            Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,
            Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,
            Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y.,
            Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
            Yoshino,M. and Hayashizaki,Y.
  TITLE     Collection, mapping, and annotation of over 28,000 cDNA clones from
            japonica rice
  JOURNAL   Science 301 (5631), 376-379 (2003)
```

MEDLINE 22752273
PUBMED 12869764
REFERENCE 2
AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K., Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K., Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M., Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M., Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A., Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakahama,Y., Nakamura,M., Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K., Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H., Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K., Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S., Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W., Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and Yoshimura,A.
TITLE Collection, mapping, and annotation of 28K full-length cDNA clones from japonica rice
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1620)
AUTHORS Kikuchi,S.
TITLE Direct Submission
JOURNAL Submitted (31-JAN-2003) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)
COMMENT This clone is one of the 32K full-length cDNA clones from japonica rice.
URL : http://cdna01.dna.affrc.go.jp/cDNA/
NIAS Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., Yamamoto,M. and Nakahama,Y.
FAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y., Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M., Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J., Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S., Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S., Yoshimura,A., Matsubara,K. and Murakami,K.
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J., Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N., Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K., Yasunishi,A. and Hayashizaki,Y.
FEATURES Location/Qualifiers
source 1. .1620
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="001-133-F04"
ORIGIN

Query Match 51.2%; Score 1032.2; DB 8; Length 1620;
Best Local Similarity 81.8%; Pred. No. 4.1e-131;
Matches 1203; Conservative 1; Mismatches 264; Indels 3; Gaps 1;

```
Qy   135 AAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATC 194
Db    31 AAGAAGAAGAACGCCAGCTATGCATTCACCTGCGCCATCCTCGCCTCCATGGCCTCCATC 90

Qy   195 CTCCTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTC 254
Db    91 ATTCTCGGCTATGACATCGGGGTGATGAGCGGCGCATCGCTGTACATCAAGAAAGACCTG 150

Qy   255 AACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATC 314
Db   151 AAGATCACTGACGTGCAGGTGGAGATCCTGATGGGCATCCTGAACATCTACTCGCTCGTC 210

Qy   315 GGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTC 374
Db   211 GGCTCCTTCGCCGCGGGGAGGACGGCGGACTGGATCGGCCGGCGGTTCACCGTCGTCTTC 270

Qy   375 GCCGCCGTCATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATG 434
Db   271 GCTGCCGCCTTCTTCTTCGCCAGAGCTCTACTCATGGGGTTCTCCGGTGACTACGCCACT 330

Qy   435 CTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCG 494
Db   331 CTCATGGTCGGCCGCTTCGTGGCGGGCGTCGGCGTGGGGTACGCGATCATGATCGCGCCC 390

Qy   495 GTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAG 554
Db   391 GTGTACACGGCCGAGATCTCGCCGGCGTCGTCGCGTGGCTTCCTCACGTCGTTCCCGGAG 450

Qy   555 GTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTG 614
Db   451 GTCTCCATCAACCTCGGTATCCTCCTCGGCTGCGTCTCCAACTATGCTTTCGCGCGCCTG 510

Qy   615 CCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTC 674
Db   511 CCGCTCTCCCTCGGGTGGCGCGTCATGCTCGGCGTCGGCGCGGCGCCATCCGTGCTGCTC 570

Qy   675 GCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTC 734
Db   571 GCGCTCATGGTGCTCGGCATGCCGGAGTCGCCACGGTGGCTTGTCATGAAGGGTCGCCTC 630

Qy   735 GCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGC 794
Db   631 GCCGACGCCAAGGCTGTGCTGGAGAAGATCGCTGACACGCCGGAGGAGGCGTCGGAGCGG 690

Qy   795 CTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACC 854
Db   691 CTCGCTGACATCAAGGCGGCGGCTGGCATCCCAGATGACCTCGACGGCGACGTGGTCACC 750

Qy   855 GTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCG 914
Db   751 GTGTCCAAGAAGAGAGGCGGCGAGGAGGGGCAGGTTTGGAGGGAGCTCGTCGTGTCCCCG 810

Qy   915 ACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCG 974
Db   811 ACCCCGGCCATGCGGCGAATCGTGCTCGCGGCGGTCGGCCTCCACTTCTTCCAGCAGGCG 870

Qy   975 TTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACG 1034
Db   871 TCGGGCGTCGACTCCGTCGTGCTCTACAGCCCGCGCGTGTTCCAGAGCGCGGGCATCACC 930

Qy  1035 AACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATC 1094
Db   931 GGCGACGACCAGCTCCTGGGCACCACCTGCGCCGTCGGCTTCGCCAAGACGGTCTTCATC 990

Qy  1095 TTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACG 1154
Db   991 CTGGTCGCCGCGTTCCTGCTCGACCGCGCCGGCCGGCGGCCGCTGCTGCTGACCAGCAGG 1050

Qy  1155 GGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCAC 1214
```

```
Db       1051 GGCGGGATGGTCTTCTCCCTCGTCGGCCTCGCGACGGGGCTCACCGTCGTGGGCGGCAGC 1110
Qy       1215 CCCGACGCCAAGATACC---TTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTAC 1271
Db       1111 CCGGACGCCCAGGTGCCCAGCTGGGCCGTCGGCCTGTGCGTCGCGTCCATCCTGGCCTAC 1170
Qy       1272 GTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTC 1331
Db       1171 GTCGCCTTCTTCTCCGTCGGCCTCGGGCCCATGTCCGGCGTGTACACCTCGGAGATCTTC 1230
Qy       1332 CCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGC 1391
Db       1231 CCGCTGCGGGCGCGCGCGCTGGGCTTCGCGGTCGCCGTGGCGTGCAACCGCGTCACCAGC 1290
Qy       1392 GGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTC 1451
Db       1291 GGCGTCATCTCCATGACCTTCCTGTCGCTCTCCAGCGCCATCACCATCGGCGGCAGCTTC 1350
Qy       1452 TTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAG 1511
Db       1351 TTCCTCTACGCCGCCATCTCGTCGCTCGCGTGGGTGTTCTTCTTCACTCGCCTCCCGGAG 1410
Qy       1512 ACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCG 1571
Db       1411 ACGCGCGGCCAGACGCTGGAGGAGATCGGCAAGGTGTTCGGCATGGACGACACGGCCATG 1470
Qy       1572 GAATCAGACGAGCCAGCCAAGGAGAAGAAGA 1602
Db       1471 GAAGCAGAGGACTCCGCCGCGTACAGGGAAA 1501

RESULT 9
AR208574
LOCUS       AR208574                1914 bp    DNA     linear   PAT 20-JUN-2002
DEFINITION  Sequence 19 from patent US 6383776.
ACCESSION   AR208574
VERSION     AR208574.1  GI:21509763
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1914)
  AUTHORS   Allen,S.M., Hitz,W.D., Kinney,A.J. and Tingey,S.V.
  TITLE     Plant sugar transport proteins
  JOURNAL   Patent: US 6383776-A 19 07-MAY-2002;
FEATURES             Location/Qualifiers
     source          1. .1914
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             49.8%;  Score 1005;  DB 6;  Length 1914;
  Best Local Similarity   79.2%;  Pred. No. 2.1e-127;
  Matches 1220;  Conservative    1;  Mismatches  311;  Indels    9;  Gaps    2;

Qy         95 GATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTT 154
Db         55 GATGGCTTCCGACGAGCTCGCAAAGGCCGTCGAGCCCAGGAAGAAGGGCAACGTCAAGTA 114
Qy        155 CGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGG 214
Db        115 TGCCTCCATATGTGCCATCCTGGCCTCCATGGCCTCTGTCATCCTTGGCTATGACATTGG 174
Qy        215 GGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGT 274
Db        175 GGTGATGAGTGGAGCGGCCATGTACATCAAGAAGGACCTGAATATCACGGACGTGCAGCT 234
Qy        275 GGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCG 334
Db        235 GGAGATCCTGATCGGGATCCTCAGTCTCTACTCGCTGTTCGGATCCTTCGCTGGCGCGCG 294
Qy        335 GACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGC 394

Db        295 GACGTCCGACAGGATCGGGCGCCGCTTGACCGTCGTGTTCGCCGCTGTCATCTTCTTCGT 354
Qy        395 GGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGT 454
Db        355 GGGCTCGTTGCTCATGGGTTTCGCCGTCAACTACGGCATGCTCATGGCGGGCCGCTTCGT 414
Qy        455 GGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTC 514
Db        415 GGCCGGAGTCGGTGTGGGCTACGGGGGCATGATCGCGCCCGTGTACACGGCCGAGATCTC 474
Qy        515 GCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCAT 574
Db        475 GCCTGCGGCGTCCCGTGGCTTCCTGACCACCTTCCCGGAGGTGTTCATCAACATCGGCAT 534
Qy        575 CCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCG 634
Db        535 CCTGCTTGGCTACCTGTCCAACTTCGCGTTCGCGCGCCTCCCGCTCCACCTCGGCTGGCG 594
Qy        635 CATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCAT 694
Db        595 CGTCATGCTCGCCATTGGCGCAGTTCCGTCCGGCCTGCTCGCGCTCCTGGTGTTCTGCAT 654
Qy        695 GCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCT 754
Db        655 GCCCGAGTCGCCTCGGTGGCTGGTCTTGAAGGGCCGCCTCGCGGACGCCAGGGCTGTGCT 714
Qy        755 GGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGC 814
Db        715 AGAGAAGACCTCTGCCACGCCAGAGGAGGCCGCCGAGCGGCTGGCCGACATCAAGGCCGC 774
Qy        815 CGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCC---CAAGAGAGGGAG 871
Db        775 GGCGGGGATTCCGAAGGGCCTCGACGGGGACGTAGTCACCGTACCCGGCAAGGAGCAAGG 834
Qy        872 CGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCG 931
Db        835 CGGCGGTGAGTTGCAGGTGTGGAAGAAGCTCATCCTGTCCCCGACCCCGGCTGTCCGACG 894
Qy        932 CATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGT 991
Db        895 CATACTGCTCTCGGCCGTGGGTCTCCACTTCTTCCAGCAGGCTTCTGGCAGCGACTCCGT 954
Qy        992 CGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTT 1051
Db        955 CGTCCAGTACAGCGCCCGCCTGTTCAAGAGCGCGGGGATCACCGACGACAACAAGCTCCT 1014
Qy       1052 GGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTT 1111
Db       1015 GGGCGTCACCTGCGCGGTGGGCGTGACCAAGACGTTCTTCATCCTGGTGGCCACGTTCCT 1074
Qy       1112 CATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTC 1171
Db       1075 GCTGGACCGCGCGGGGCGTCGGCCTCTGCTGCTGATCAGCACGGGCGGGATGATTGTCTC 1134
Qy       1172 CCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACC 1231
Db       1135 GCTCATCTGCCTCGGGTCGGGGCTCACCGTCGCGGGGCATCACCCGGACACCAAGGTCGC 1194
Qy       1232 TTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGG 1291
Db       1195 GTGGGCCGTCGCCCTGTGCATCGCGTCAACCCTGTCCTACATCGCCTTCTTCTCCATCGG 1254
Qy       1292 CCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCT 1351
Db       1255 CCTCGGGCCCATCACGGGCGTGTACACCTCGGAAATATTCCCGCTGCAGGTGCGCGCGCT 1314
Qy       1352 GGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTT 1411
Db       1315 GGGCTTCGCGGTGGGTGTGGCGAGCAACCGCGTCACCAGCGCCGTCATCTCCATGACCTT 1374
Qy       1412 CCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGC 1471
```

```
Db       1375 CCTGTCCCTCTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGC 1434

Qy       1472 CGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGA 1531

Db       1435 CGCGGTCGCTTGGGTTTTCTTCTTCACGTGCCTCCCGGAGACACGCGGCCGGACGCTGGA 1494

Qy       1532 GGAGATGAGCAAGCTGTTCGGC------GACACGGCCGCCGCCTCGGAATCAGACGAGCC 1585

Db       1495 GGAGATGGGCAAGCTGTTCGGCATGCCAGACACGGGCATGGCTGAAGAAGCAGAAGACGC 1554

Qy       1586 AGCCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCACTAACT 1626

Db       1555 CGCAGCCAAGGAGAAGGTGGTGGAACTGCCTAGCAGCAAGT 1595
```

RESULT 10
BT009301
LOCUS BT009301 1872 bp mRNA linear PLN 20-JUN-2003
DEFINITION Triticum aestivum clone wlm1.pk0012.h1:fis, full insert mRNA sequence.
ACCESSION BT009301
VERSION BT009301.1 GI:32128852
KEYWORDS FLI_CDNA.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 1872)
AUTHORS Tingey,S.V., Wolters,P., Powell,W., Dolan,M., Miao,G.-H., Caraher,N.R., Hanafey,M.K. and Hainey,C.F.
TITLE Direct Submission
JOURNAL Submitted (20-JUN-2003) Crop Genetics, E. I. DuPont de Nemours and Company, 1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
FEATURES Location/Qualifiers
source 1. .1872
/organism="Triticum aestivum"
/mol_type="mRNA"
/db_xref="taxon:4565"
/clone="wlm1.pk0012.h1:fis"
ORIGIN

Query Match 49.0%; Score 988; DB 8; Length 1872;
Best Local Similarity 77.7%; Pred. No. 4.2e-125;
Matches 1192; Conservative 1; Mismatches 341; Indels 0; Gaps 0;

```
Qy         99 GCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCC 158

Db        128 GCTGCGCTCCCGGAGCCGGGGGCAGTCCATCCAAGGAACAAGGGCAATTTCAAGTACGCC 187

Qy        159 TTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTG 218

Db        188 TTCACCTGCGCCCTCTGTGCTTCCATGGCCACCATCGTCCTCGGCTACGACGTTGGGGTG 247

Qy        219 ATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAG 278

Db        248 ATGAGCGGTGCGTCGCTGTACATCAAGAGGGACCTGCAGATCACGGACGTGCAGCTGGAG 307

Qy        279 GTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACG 338

Db        308 ATCATGATGGGCATCCTGAGCGTGTACGCGCTCATCGGGTCCTTCCTCGGCGCGAGGACG 367

Qy        339 TCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGG 398

Db        368 TCCGACTGGGTCGGCCGCCGCGTCACCGTCGTCTTCGCGGCCGCCATCTTCACCACCGGC 427

Qy        399 GSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCC 458

Db        428 TCCTTGCTCATGGGCTTCGCGGTCAACTACGCCATGCTCATGGTCGGCCGCTTCGTCACC 487

Qy        459 GGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCG 518

Db        488 GGCATCGGCGTGGGCTACGCCATCATGGTCGCGCCAGTGTACACGGCCGAGGTGTCCCCG 547

Qy        519 GCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTG 578

Db        548 GCGTCGGCCCGCGGCTTCCTCACGTCTTTCACCGAGGTGTTCATCAATGTGGGCATCCTC 607

Qy        579 CTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATC 638

Db        608 CTTGGCTACGTCTCCAACTACGCCTTCGCGCGCCTCCCGCTCCACCTCAGCTGGCGCGTC 667

Qy        639 ATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCG 698

Db        668 ATGCTCGGCATCGGCGCCGTCCCGTCCGCCCTGCTTGCGCTCATGGTGTTCGGCATGCCG 727

Qy        699 GAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAG 758

Db        728 GAGTCTCCTCGCTGGCTCGTCATGAAAGGCCGCCTCGCGGACGCCAGGGCCGTTCTGGCC 787

Qy        759 AAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCC 818

Db        788 AAGACCTCCGACACGCCGGAGGAGGCCGTGGAGCGCCTTGACCAGATCAAGGCTGCCGCC 847

Qy        819 GGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAAC 878

Db        848 GGCATCCCTAGGGAACTTGACGGCGACGTGGTCGTCATGCCTAAGACAAAAGGCGGCCAG 907

Qy        879 GAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTG 938

Db        908 GAGAAGCAGGTGTGGAAGGAGCTCATCTTTTCGCCGACCCCAGCCATGCGGCGCATACTG 967

Qy        939 CTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTC 998

Db        968 CTCGCGGCGCTCGGCATCCATTTCTTTCAGCAGGCGACGGGCTCCGACTCCGTCGTGCTC 1027

Qy        999 TACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGCACC 1058

Db       1028 TATAGCCCACGCGTGTTCCAGAGCGCGGGCATCACCGGCGACAACCACCTGCTCGGCGCC 1087

Qy       1059 ACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGAC 1118

Db       1088 ACATGCGCCATGGGGGTCATGAAGACGCTCTTCATCCTGGTGGCCACGTTCCAGCTCGAC 1147

Qy       1119 GGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATC 1178

Db       1148 CGCGTCGGCAGGCGGCCGCTGCTGCTGACCAGCACGGCCGGCATGCTCGCCTGTCTCATC 1207

Qy       1179 GGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCC 1238

Db       1208 GGCCTCGGGACGGGCCTCACCGTCGTGGGTCGGCACCCGGACGCCAAGGTCCCGTGGGCC 1267

Qy       1239 ATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGC 1298

Db       1268 ATCGGCCTGTGCATCGTGTCCATCTTGGCCTACGTGTCCTTCTTCTCCATCGGCCTCGGG 1327

Qy       1299 CCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGC 1358

Db       1328 CCCCTCACCAGCGTGTACACCTCGGAGGTCTTCCCACTGCGGGTGCGCGCGCTGGGCTTC 1387

Qy       1359 TCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCG 1418

Db       1388 GCGCTGGGCACGTCATGCAACCGCGTCACCAGCGCCGCGGTCTCCATGTCCTTCCTGTCC 1447

Qy       1419 CTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTC 1478

Db       1448 TTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTGTACGCCGGCATCGCGGCGATA 1507

Qy       1479 GCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATG 1538

Db       1508 GGATGGATTTTCTTCTTCACCTTCATTCCGGAGACGCGTGGCCTGCCGCTCGAGGAGATA 1567

Qy       1539 AGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAGCCAAGGAGAAG 1598

Db       1568 GGGAAGCTTTTCGGCATGACGGACACGGCCGTCGAAGCCCAAGACACCGCCACGAAAGAC 1627
```

```
Qy       1599 AAGAAGGTGGAAATGGCCGCCACTAACTGATCAA 1632
              |||  |   | |  ||     |   | ||| | |
Db       1628 AAGGCGAAAGTAGGGGAGATGAACTAGTGAGCTA 1661


RESULT 11
AP004185/c
LOCUS       AP004185               25270 bp    DNA     linear   PLN 02-SEP-2003
DEFINITION  Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 7,
            BAC clone:OJ1301_C12.
ACCESSION   AP004185
VERSION     AP004185.3  GI:32261236
KEYWORDS    .
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1
  AUTHORS   Sasaki,T., Matsumoto,T. and Yamamoto,K.
  TITLE     Oryza sativa nipponbare(GA3) genomic DNA, chromosome 7, BAC
            clone:OJ1301_C12
  JOURNAL   Published Only in Database (2001)
REFERENCE   2  (bases 1 to 25270)
  AUTHORS   Sasaki,T., Matsumoto,T. and Yamamoto,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-SEP-2001) Takuji Sasaki, National Institute of
            Agrobiological Sciences, Rice Genome Research Program; Kannondai
            2-1-2, Tsukuba, Ibaraki 305-8602, Japan
            (E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
            Tel:81-298-38-7441, Fax:81-298-38-7468)
COMMENT     On Jun 25, 2003 this sequence version replaced gi:21280394.
            Genes were predicted from the integrated results of the following:
            GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), FGENESH
            (http://www.softberry.com/), GeneMark.hmm
            (http://opal.biology.gatech.edu/GeneMark/), GlimmerM
            (http://www.tigr.org/tdb/glimmerm/glmr_form.html), RiceHMM
            (http://rgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor
            (http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4
            (http://globin.cse.psu.edu/html/docs/sim4.html), gap2
            (http://www.tigr.org/software/glimmerm/), BLASTN and BLASTX. The
            genomic sequence was searched against NCBI NonRedundant Protein
            database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA
            sequence database at RGP or DDBJ. Protein homologies of the coding
            regions were searched against NCBI NonRedundant Protein database
            with BLASTP. ESTs represent the identified cDNA sequences using
            BLASTN with the corresponding DDBJ accession no. and RGP clone ID.
            Full-length cDNAs represent the identified cDNA sequences using
            BLASTN with the corresponding DDBJ accession no.
            A gene with identity or significant homology to a protein is
            classified based on the protein name to indicate the homology level
            such as same name, 'putative-' and '-like protein'.  A gene without
            significant homology to any protein but with full-length cDNA or
            EST homology (covering almost the entire length of partial
            sequence) is classified as an 'unknown' protein. A gene predicted
            by two or more gene prediction programs is classified as a
            'hypothetical' protein according to IRGSP standard. A gene
            predicted by a single gene prediction program is also classified as
            a probable 'hypothetical' protein and is included as a
            miscellaneous feature of the sequence.
            The orientation of the sequence is from -21M13 to M13rev of the BAC
            clone. This is a partial sequence of OJ1301_C12 clone.This sequence
            of OJ1301_C12 clone has an overlap with P0453G03 clone(DDBJ:
            AP004276)at 5' end and an overlap with OJ1127_E01(DDBJ: AP003747)
            at 3' end.The sequence was generated by combining Monsanto and
            RGP-Japan sequencing data. Detailed information on overlap and
            assembly quality together with annotation of this entry is
            available at
            http://rgp.dna.affrc.go.jp/GenomeSeq.html.
FEATURES             Location/Qualifiers
     source          1. .25270
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="genomic DNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /chromosome="7"
                     /clone="OJ1301_C12"
     gene            779. .1103
                     /gene="OJ1301_C12.1"
     mRNA            join(<779. .940,1011. .>1103)
                     /gene="OJ1301_C12.1"
                     /note="start and end point are not identified"
     CDS             join(779. .940,1011. .1103)
                     /gene="OJ1301_C12.1"
                     /note="predicted by GENSCAN etc."
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="BAC83309.1"
                     /db_xref="GI:34393629"
                     /translation="MYRERAQRRGRSSATVFLASSCRQLGKGKKVFARDGNPAGCGYP
                     CRPASARICTLIRLSADWTTAACDDGSSAAVDFHGSFRFS"
     gene            complement(join(1344. .1480,2498. .2686,3144. .3198))
                     /gene="OJ1301_C12.2"
     misc_feature    complement(join(1344. .1480,2498. .2686,3144. .3198))
                     /gene="OJ1301_C12.2"
                     /note="hypothetical ORF
                     predicted by GENSCAN
                     this category is not included in IRGSP standard"
     gene            5120. .7180
                     /gene="OJ1301_C12.3"
     mRNA            join(5120. .5308,5408. .7180)
                     /gene="OJ1301_C12.3"
                     /note="supported by full-length cDNA(s): AK059423"
     mRNA            join(5121. .5308,5408. .7180)
                     /gene="OJ1301_C12.3"
                     /note="supported by full-length cDNA(s): AK102676"
     CDS             join(5197. .5308,5408. .6828)
                     /gene="OJ1301_C12.3"
                     /note="contains EST(s):
                     C20132(E11563),AU163952(E11563),AU097469(S5067),
                     AU033198(S5067)
                     contains full-length cDNA(s): AK059423,AK102676"
                     /codon_start=1
                     /product="putative sorbitol transporter"
                     /protein_id="BAC83310.1"
                     /db_xref="GI:34393630"
                     /translation="MASAALPEAVAPKKKGNVRFAFACAILASMTSILLGYDIGVMSG
                     ASLYIKKDFNISDGKVEVLMGILNLYSLIGSFAAGRTSDWIGRRYTIVFAAVIFFAGA
                     FLMGFAVNYAMLMFGRFVAGIGVGYALMIAPVYTAEVSPASARGFLTSFPEVFINFGI
                     LLGYVSNYAFSRLPLNLGWRIMLGIGAAPSVLLALMVLGMPESPRWLVMKGRLADAKV
                     VLEKTSDTAEEAAERLADIKAAAGIPEELDGDVVTVPKRGSGNEKRVWKELILSPTPA
                     MRRILLSGIGIHFFQQASGIDSVVLYSPRVFKSAGITDDKHLLGTTCAVGVTKTLFIL
                     VATFFLDRVGRRPLLLSSTGGMILSLIGLGAGLTVVGQHPDAKIPWAIGLSIASTLAY
                     VAFFSIGLGPITWVYSSEIFPLQVRALGCSLGVAANRVTSGVISMTFLSLSKAITIGG
                     SFFLYSGIAALAWVFFYTYLPETRGRTLEEMSKLFGDTAAASESDEPAKEKKKVEMAA
                     TN"
     gene            complement(7647. .9324)
                     /gene="OJ1301_C12.4"
     mRNA            complement(join(<7647. .9067,9210. .>9324))
                     /gene="OJ1301_C12.4"
                     /note="start and end point are not identified"
     CDS             complement(join(7647. .9067,9210. .9324))
                     /gene="OJ1301_C12.4"
                     /codon_start=1
                     /product="putative sorbitol transporter"
                     /protein_id="BAC83311.1"
                     /db_xref="GI:34393631"
                     /translation="MPESTKLTAMAVDPKKKNASYAFTCAILASMASIILGYDIGVMS
                     GASLYIKKDLKITDVQVEILMGILNIYSLVGSFAAGRTADWIGRRFTVVFAAAFFFAR
                     ALLMGFSGDYATLMVGRFVAGVGVGYAIMIAPVYTAEISPASSRGFLTSFPEVSINLG
                     ILLGYVSNYAFARLPLSLGWRVMLGVGAAPSVLLALMVLGMPESPRWLVMKGRLADAK
                     AVLEKIADTPEEASERLADIKAAAGIPDDLDGDVVTVSKKRGGEEGQVWRELVVSPTP
                     AMRRIVLAAVGLHFFQQASGVDSVVLYSPRVFQSAGITGDDQLLGTTCAVGFAKTVFI
                     LVAAFLLDRAGRRPLLLTSTGGMVFSLVGLATGLTVVGGSPDAQVPSWAVGLCVASIL
                     AYVAFFSVGLGPMSGVYTSEIFPLRARALGFAVAVACNRVTSGVISMTFLSLSSAITI
```

```
                GGSFFLYAAISSLAWVFFFTRLPETRGQTLEEIGKVFGMDDTAMEAEDSAAYRERLLA
                TSP"
gene            11758. .12731
                /gene="OJ1301_C12.5"
mRNA            join(<11758. .11894,11990. .12071,12243. .>12731)
                /gene="OJ1301_C12.5"
                /note="start and end point are not identified"
CDS             join(11758. .11894,11990. .12071,12243. .12731)
                /gene="OJ1301_C12.5"
                /note="predicted by FGENESH etc."
                /codon_start=1
                /product="hypothetical protein"
                /protein_id="BAC83312.1"
                /db_xref="GI:34393632"
                /translation="MAKRARTLTGDHLRSSSSSNQAPSWGAKRARSGSALRDRSLPDS
                EWRDWSSIGNGPAWLISERVLAGDDVADYADLPELRGHRVFGPTAEGLLVLLDEATFA
                VRLLNPLTRQLTDLPPATTVLDPTSLMISPVHYLLHVLTVHGAGLADGSATVVVYFEN
                ARKLAAAKPGDERWTVVEPDASFASVSSFAGRFYCATRHAIRPVIVEVRTTDHRLPPQ
                LVVVAKLAGVPSFPRNC"
gene            join(12995. .17652,18451. .19187)
                /gene="OJ1301_C12.6"
repeat_unit     12995. .13731
                /gene="OJ1301_C12.6"
                /note="5' Terminal repeat"
                /rpt_type=terminal
misc_feature    13837. .17652
                /gene="OJ1301_C12.6"
                /note="probably inactive due to including stop codon(s) in
                CDS
                pseudogene,Fourf gag/pol protein"
                /pseudo
repeat_unit     18451. .19187
                /gene="OJ1301_C12.6"
                /note="3' Terminal repeat"
                /rpt_type=terminal
gene            complement(19137. .20909)
                /gene="OJ1301_C12.7"
mRNA            complement(19137. .20909)
                /gene="OJ1301_C12.7"
                /note="supported by full-length cDNA(s): AK110773"
gene            19581. .21159
                /gene="OJ1301_C12.8"
mRNA            19581. .21159
                /gene="OJ1301_C12.8"
                /note="supported by full-length cDNA(s): AK069022"
CDS             complement(19794. .20309)
                /gene="OJ1301_C12.7"
                /note="contains full-length cDNA(s): AK110773"
                /codon_start=1
                /product="unknown protein"
                /protein_id="BAC83313.1"
                /db_xref="GI:34393633"
                /translation="MPSSAPASQRSHRAHAGLTASKPPPTARSFRWPHLPPCRPRQPP
                QPPSAPPRQPTLPPGVSPLTARASSRAPCRPPPPPAVPATCCLLPRLPPPASCYHLRR
                CELRRSETCLAMTSASSASRRTAPRVPPPPESPAPDPATGRPDQPPRGPNPPPAWPDS
                PCGPLRTPSSA"
CDS             20013. .20480
                /gene="OJ1301_C12.8"
                /note="contains EST(s):
                D40430(S2406),D15710(C1130A),AU086036(S2406)
                contains full-length cDNA(s): AK069022"
                /codon_start=1

Query Match             48.9%;  Score 985.8;  DB 8;  Length 25270;
Best Local Similarity   82.0%;  Pred. No. 7.7e-125;
Matches 1147;  Conservative    1;  Mismatches  248;  Indels    3;  Gaps    1;

Qy       207 GATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGAC 266
             || |||||||||||||||||| || |||||||||||||||| ||| | || |||| ||||
Db      9068 GACATCGGGGTGATGAGCGGCGCATCGCTGTACATCAAGAAAGACCTGAAGATCACTGAC 9009

Qy       267 GGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCG 326
             | | |||||||| | || |||||||| |||||| |||||||||| ||||||||||||||
```

```
Db      9008 GTGCAGGTGGAGATCCTGATGGGCATCCTGAACATCTACTCGCTCGTCGGCTCCTTCGCC 8949

Qy       327 GCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATA 386
             |||||| ||||| |||||||||||||||||||||| |||| |||| ||||| |||| | |
Db      8948 GCGGGGAGGACGGCGGACTGGATCGGCCGGCGGTTCACCGTCGTCTTCGCTGCCGCCTTC 8889

Qy       387 TTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGC 446
             |||||||| | |:  | |||||||||||| ||| |||||||| |||||| |||||
Db      8888 TTCTTCGCCAGAGCTCTACTCATGGGGTTCTCCGGTGACTACGCCACTCTCATGGTCGGC 8829

Qy       447 CGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCC 506
             |||||||||||| ||| |||||||||| |||||| |||||||||||||| |||||||| |||
Db      8828 CGCTTCGTGGCGGGCGTCGGCGTGGGGTACGCGATCATGATCGCGCCCGTGTACACGGCC 8769

Qy       507 GAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAAC 566
             ||| | |||||||||||| |||||||||||| |||||||||||||||| | |||||||
Db      8768 GAGATCTCGCCGGCGTCGTCGCGTGGCTTCCTCACGTCGTTCCCGGAGGTCTCCATCAAC 8709

Qy       567 TTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTC 626
             |||| |||| ||||| |||||||| |||||||||||| | ||| |||||||   ||||
Db      8708 CTCGGTATCCTCCTCGGCTACGTCTCCAACTATGCTTTCGCGCGCCTGCCGCTCTCCCTC 8649

Qy       627 GGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTG 686
             ||||||||| |||||||||| |||||||||||| |||||||||||||||||||||||||||
Db      8648 GGGTGGCGCGTCATGCTCGGCGTCGGCGCGGCGCCATCCGTGCTGCTCGCGCTCATGGTG 8589

Qy       687 CTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAG 746
             |||||||||||||||||||| |||||||| |||||||||| |||||||| |||||||||
Db      8588 CTCGGCATGCCGGAGTCGCCACGGTGGCTTGTCATGAAGGGTCGCCTCGCCGACGCCAAG 8529

Qy       747 GTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATC 806
             | |||||||||||||| | | |||||| |||||||||  ||||||| || || ||||||
Db      8528 GCTGTGCTGGAGAAGATCGCTGACACGCCGGAGGAGGCGTCGGAGCGGCTCGCTGACATC 8469

Qy       807 AAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGA 866
             ||||| || || |||||||| || || |||||||||||||||| |||||  ||||||
Db      8468 AAGGCGGCGGCTGGCATCCCAGATGACCTCGACGGCGACGTGGTCACCGTGTCCAAGAAG 8409

Qy       867 GGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATG 926
             | ||||  | |||  || ||| |||| |||||||  || ||||||||||||||||||||
Db      8408 AGAGGCGGCGAGGAGGGGCAGGTTTGGAGGGAGCTCGTCGTGTCCCCGACCCCGGCCATG 8349

Qy       927 CGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCAC 986
             ||||| ||| ||||  | | | ||||| |||||||||||||||| |||| ||||  |  ||
Db      8348 CGGCGAATCGTGCTCGCGGCGGTCGGCCTCCACTTCTTCCAGCAGGCGTCGGGCGTCGAC 8289

Qy       987 TCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACAC 1046
             ||||||||  |||||||||| | ||||||| ||||| | ||  | ||   |||| | ||
Db      8288 TCCGTCGTGCTCTACAGCCCGCGCGTGTTCCAGAGCGCGGGCATCACCGGCGACGACCAG 8229

Qy       1047 TTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACT 1106
              || |||||||||| ||  |  ||||  || |||||| | | ||||||  || | ||  |
Db      8228 CTCCTGGGCACCACCTGCGCCGTCGGCTTCGCCAAGACGGTCTTCATCCTGGTCGCCGCG 8169

Qy       1107 TTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATC 1166
              ||| |  ||||| ||| ||| |||||||||||| |||||  ||||||||||||||||  ||
Db      8168 TTCCTGCTCGACCGCGCCGGCCGGCGGCCGCTGCTGCTGACCAGCACGGGCGGGATGGTC 8109

Qy       1167 CTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAG 1226
              |||||||| |||||||||   | |||||||||||||| |||   ||| |||||| ||
Db      8108 TTCTCCCTCGTCGGCCTCGCGACGGGGCTCACCGTCGTGGGCGGCAGCCCGGACGCCCAG 8049

Qy       1227 ATACC---TTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTC 1283
              | ||    |||||| |||||||  || |||| |||| ||| ||||||||||||||||||
Db      8048 GTGCCCAGCTGGGCCGTCGGCCTGTGCGTCGCGTCCATCCTGGCCTACGTCGCCTTCTTC 7989

Qy       1284 TCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTG 1343
              ||| ||||||| || |||||  |  | ||||||| ||||||||||||||||||| | || |
Db      7988 TCCGTCGGCCTCGGGCCCATGTCCGGCGTGTACACCTCGGAGATCTTCCCGCTGCGGGCG 7929

Qy       1344 CGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCC 1403
              |||||||||||| | || ||| ||| ||   ||||||||||||||||||||||||||||
Db      7928 CGCGCGCTGGGCTTCGCGGTCGCCGTGGCGTGCAACCGCGTCACCAGCGGCGTCATCTCC 7869
```

```
Qy       1404 ATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCC 1463
             |||||||||||||||||| ||||  ||||||||||||||||||||||||||||||||||| ||
Db       7868 ATGACCTTCCTGTCGCTCTCCAGCGCCATCACCATCGGCGGCAGCTTCTTCCTCTACGCC 7809

Qy       1464 GGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGG 1523
             | |||| |  ||||||| |||||||||||| |||   |||||||||||| ||||||| |
Db       7808 GCCATCTCGTCGCTCGCGTGGGTGTTCTTCTTCACTCGCCTCCCGGAGACGCGCGGCCAG 7749

Qy       1524 ACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAG 1583
             |||||||||||||   ||||| ||||||||      | | | |||  |||| |||| ||
Db       7748 ACGCTGGAGGAGATCGGCAAGGTGTTCGGCATGGACGACACGGCCATGGAAGCAGAGGAC 7689

Qy       1584 CCAGCCAAGGAGAAGAAGA 1602
              | |||  | | | | | |
Db       7688 TCCGCCGCGTACAGGGAAA 7670


RESULT 12
AR208578
LOCUS       AR208578               1872 bp    DNA     linear   PAT 20-JUN-2002
DEFINITION  Sequence 27 from patent US 6383776.
ACCESSION   AR208578
VERSION     AR208578.1  GI:21509768
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1872)
  AUTHORS   Allen,S.M., Hitz,W.D., Kinney,A.J. and Tingey,S.V.
  TITLE     Plant sugar transport proteins
  JOURNAL   Patent: US 6383776-A 27 07-MAY-2002;
FEATURES             Location/Qualifiers
     source          1. .1872
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             48.7%;  Score 981.6;  DB 6;  Length 1872;
  Best Local Similarity   77.4%;  Pred. No. 3.2e-124;
  Matches 1188;  Conservative    1;  Mismatches  345;  Indels    0;  Gaps    0;

Qy         99 GCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCC 158
              ||| |   | ||  ||||| ||| |||   || | ||| ||||||||  ||  || ||||
Db        128 GCTGCGCTCCCGGAGCCGGGGGCAGTCCATCCAAGGAACAAGGGCAATTTCAAGTACGCC 187

Qy        159 TTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTG 218
              ||| |||||||| ||   || |||||| || ||||| |||||||||||||  | ||||||
Db        188 TTCACCTGCGCCCTCTGTGCTTCCATGGCCACCATCGTCCTCGGCTACGACGTTGGGGTG 247

Qy        219 ATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAG 278
              |||||||| ||||||||||||||||||| |||| |  | ||||  |||| | || |||||
Db        248 ATGAGCGGTGCGTCGCTGTACATCAAGAGGGACCTGCAGATCACGGACGTGCAGCTGGAG 307

Qy        279 GTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACG 338
              |  | |||||||| |||| | | ||| |||||||||| ||||||   |  | | |||||
Db        308 ATCATGATGGGCATCCTGAGCGTGTACGCGCTCATCGGGTCCTTCCTCGGCGCGAGGACG 367

Qy        339 TCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGG 398
              || |||||| ||||||| ||    |||| |||| ||||| |||| ||| |||  |   ||
Db        368 TCCGACTGGGTCGGCCGCCGCGTCACCGTCGTCTTCGCGGCCGCCATCTTCAACAACGGC 427

Qy        399 GSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCC 458
              : || |||||||| ||||| |||||||||||||||||||| |||| ||||||||  ||
Db        428 TCCTTGCTCATGGGCTTCGCGGTCAACTACGCCATGCTCATGGTCGGGCGCTTCGTCACC 487

Qy        459 GGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCG 518
              || |||||||||||||||  ||||| |||||||  |||||||||  |||||||||| |||
Db        488 GGAATCGGCGTGGGCTACGCCATCATGGTCGCGCCAGTGTACACGCCCGAGGTGTCCCCG 547

Qy        519 GCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTG 578
              |||||||| || |||||||| ||||| ||| | ||||||||||||||  | |||||||||
Db        548 GCGTCGGCCCGCGGCTTCCTCACGTCTTTCACCGAGGTGTTCATCAATGTGGGCATCCTC 607

Qy        579 CTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATC 638
              || || |||||||| ||||| || ||| | ||| | |||||  ||||| | |||||| ||
Db        608 CTTGGCTACGTCTCCAACTACGCCTTCGCGCGCCTCCCGCTCCACCTCAGCTGGCGCGTC 667

Qy        639 ATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCG 698
              |||||||||||||||||| |  |||||||  ||||| |||||||||||| ||||||||||
Db        668 ATGCTCGGCATCGGCGCCGTCCCGTCCGCCCTGCTTGCGCTCATGGTGTTCGGCATGCCG 727

Qy        699 GAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAG 758
              ||||| || || ||||| |||||||| || ||||||||||||||| ||  || ||||
Db        728 GAGTCTCCTCGCTGGCTCGTCATGAAAGGCCGCCTCGCGGACGCCAGGGCCGTTCTGGCC 787

Qy        759 AAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCC 818
              |||||||||||||||| |||||||||| |||||||||  | | | ||||||| ||||||
Db        788 AAGACCTCCGACACGCCGGAGGAGGCCGTGGAGCGCCTTGACCAGATCAAGGCTGCCGCC 847

Qy        819 GGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAAC 878
              ||||||||  ||| || |||||||||||||  | | || |||| |    ||||  |
Db        848 GGCATCCCTAGGGAACTTGACGGCGACGTGGTCGTCATGCCTAAGACAAAAGGCGGCCAG 907

Qy        879 GAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTG 938
              ||||||| |||||||||||||||||| | || |||||||| |||||||||||||| |||
Db        908 GAGAAGCAGGTGTGGAAGGAGCTCATCTTTTCGCCGACCCCAGCCATGCGGCGCATACTG 967

Qy        939 CTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTC 998
              ||  | | | |||||||||| ||||| ||||| |||  ||||    ||||||||||  ||
Db        968 CTCGCGGCGCTCGGCATCCATTTCTTTCAGCAGGCGACGGGCTCCGACTCCGTCGTGCTC 1027

Qy        999 TACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGCACC 1058
              || ||||| | ||||||| ||||| | ||  | ||   ||||||  |||  |  | ||| ||
Db       1028 TATAGCCCACGCGTGTTCCAGAGCGCGGGCATCACCGGCGACAACCACCTGCTCGGCGCC 1087

Qy       1059 ACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGAC 1118
              || ||  |  | || |||| |||| ||| |||||| || |||| || |||    |||||
Db       1088 ACATGCGCCATGGGGGTCATGAAGACGCTCTTCATCCTGGTGGCCACGTTCCAGCTCGAC 1147

Qy       1119 GGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATC 1178
              |||||||  |||||||||| |||||  |||||||| ||| ||  ||  ||  ||||||
Db       1148 CGCGTCGGCAGGCGGCCGCTGCTGCTGACCAGCACGGCCGGCATGCTCGCCTGTCTCATC 1207

Qy       1179 GGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCC 1238
              |||||||  | || |||||||||| || | ||||||  |||||||| | || ||||||
Db       1208 GGCCTCGGGACGGGCCTCACCGTCGTGGGTCGGCACCCGGACGCCAAGGTCCCGTGGGCC 1267

Qy       1239 ATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGC 1298
              |||||||  ||||||  |||| | | ||||||||  |||||||||||||||||| ||
Db       1268 ATCGGCCTGTGCATCGTGTCCATCTTGGCCTACGTGTCCTTCTTCTCCATCGGCCTCGGG 1327

Qy       1299 CCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGC 1358
              ||| |||| | ||||||| ||||||| ||||||| || | ||||||||||||||| |
Db       1328 CCCCTCACCAGCGTGTACACCTCGGAGGTCTTCCCACTGCGGGTGCGCGCGCTGGGCTTC 1387

Qy       1359 TCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCG 1418
              |||| |||    | |||||||||||||||| ||   |||||||| |||||||||
Db       1388 GCGCTGGGCACGTCATGCAACCGCGTCACCAGCGCCGCGGTCTCCATGTCCTTCCTGTCC 1447

Qy       1419 CTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTC 1478
              |||||||||||||||||||||||||||||||||||| ||| |||||||||| ||| |
Db       1448 TTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTGTACGCCGGCATCGCGGCGATA 1507

Qy       1479 GCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATG 1538
              | ||| | |||||| ||||| | |||||||| || |||| | |||| ||||||||
Db       1508 GGATGGATTTTCTTCTTCACCTTCATTCCGGAGACGCGTGGCCTGCCGCTCGAGGAGATA 1567

Qy       1539 AGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAGCCAAGGAGAAG 1598
              | ||||| ||||| |||| | | ||| ||| | || | |||| | | |
Db       1568 GGGAAGCTTTTCGGCATGACGGACACGGCCGTCGAAGCCCAAGACACCGCCACGAAAGAC 1627

Qy       1599 AAGAAGGTGGAAATGGCCGCCACTAACTGATCAA 1632
              |||  |  | |  ||    |  | ||| | |
Db       1628 AAGGCGAAAGTAGGGGAGATGAACTAGTGAGCTA 1661
```

```
RESULT 13
AP004365/c
LOCUS       AP004365               193577 bp    DNA     linear   PLN 27-NOV-2003
DEFINITION  Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1,
            PAC clone:P0458E05.
ACCESSION   AP004365 BA000010
VERSION     AP004365.3  GI:21902053
KEYWORDS    .
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1
  AUTHORS   Sasaki,T., Matsumoto,T., Yamamoto,K., Sakata,K., Baba,T.,
            Katayose,Y., Wu,J., Niimura,Y., Cheng,Z., Nagamura,Y.,
            Antonio,B.A., Kanamori,H., Hosokawa,S., Masukawa,M., Arikawa,K.,
            Chiden,Y., Hayashi,M., Okamoto,M., Ando,T., Aoki,H., Arita,K.,
            Hamada,M., Harada,C., Hijishita,S., Honda,M., Ichikawa,Y.,
            Idonuma,A., Iijima,M., Ikeda,M., Ikeno,M., Itoh,S., Itoh,T.,
            Itoh,Y., Itoh,Y., Iwabuchi,A., Kamiya,K., Karasawa,W., Katagiri,S.,
            Kikuta,A., Kobayashi,N., Kono,I., Machita,K., Maehara,T.,
            Mizuno,H., Mizubayashi,T., Mukai,Y., Nagasaki,H., Nakashima,M.,
            Nakama,Y., Nakamichi,Y., Nakamura,M., Namiki,N., Negishi,M.,
            Ohta,I., Ono,N., Saji,S., Sakai,K., Shibata,M., Shimokawa,T.,
            Shomura,A., Song,J., Takazaki,Y., Terasawa,K., Tsuji,K., Waki,K.,
            Yamagata,H., Yamane,H., Yoshiki,S., Yoshihara,R., Yukawa,K.,
            Zhong,H., Iwama,H., Endo,T., Ito,H., Hahn,J.H., Kim,H.I., Eun,M.Y.,
            Yano,M., Jiang,J. and Gojobori,T.
  TITLE     The genome sequence and structure of rice chromosome 1
  JOURNAL   Nature 420 (6913), 312-316 (2002)
  MEDLINE   22337376
   PUBMED   12447438
REFERENCE   2  (bases 1 to 193577)
  AUTHORS   Sasaki,T., Matsumoto,T. and Yamamoto,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-NOV-2001) Takuji Sasaki, National Institute of
            Agrobiological Sciences, Rice Genome Research Program; Kannondai
            2-1-2, Tsukuba, Ibaraki 305-8602, Japan
            (E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
            Tel:81-298-38-7441, Fax:81-298-38-7468)
COMMENT     On Jul 17, 2002 this sequence version replaced gi:19773518.
            Genes were predicted from the integrated results of the following:
            GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as SlicePredictor
            (October 1998 version). The genomic sequence was searched against
            NCBI NonRedundant Protein database, nr
            (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at
            RGP. Protein homologies of the coding regions were searched against
            NCBI NonRedundant Protein database with BLASTP2.0. ESTs represent
            the identified cDNA sequences using BLASTN 2.0 with the
            corresponding DDBJ accession no. and RGP clone ID.
            A gene with identity or significant homology to a protein is
            classified based on the protein name to indicate the homology level
            such as same name, 'putative-' and '-like protein'.  A gene without
            significant homology to any protein but with EST homology (covering
            almost the entire length of partial sequence) is classified as an
            'unknown' protein. A gene predicted with a gene prediction program
            is classified as a 'hypothetical' protein.
            The orientation of the sequence is from T7 to SP6 of the PAC clone.
            This sequence of P0458E05 clone has an overlap with P0483G10 (DDBJ:
            AP003263) at the position 1 to 50,738 of 5' end and an overlap with
            OJ1656_A11 (DDBJ: AP003448) at the position 188,119 to 193,577 of
            3' end.  Detailed information on overlap and assembly quality
            together with annotation of this entry is available at
            http://rgp.dna.affrc.go.jp/GenomeSeq.html.
FEATURES             Location/Qualifiers
     source          1. .193577
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="genomic DNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /chromosome="1"
                     /clone="P0458E05"
     gene            join(1748. .1991,2570. .2750,2997. .3243,5685. .5813,
                     5922. .6140,6199. .6213)
                     /gene="P0458E05.1"
     CDS             join(1748. .1991,2570. .2750,2997. .3243,5685. .5813,
                     5922. .6140,6199. .6213)
                     /gene="P0458E05.1"
                     /note="hypothetical protein"
                     /codon_start=1
                     /protein_id="BAC05602.1"
                     /db_xref="GI:21902054"
                     /translation="MRRRLFRRPLPRPWGAYARRSRRWRVPAVTAWYGHARRAEVCRG
                     ARWGQKKRWEEGGGGRARGFGFHGEVSSVWVVVVVVATAHSVSPRVRAADGRACPPPSL
                     TGGAHPSSPSSRRSRRPARAERRRRRRRIRRIPPPPRAIKIELAPSSSLSSVADDLRP
                     ALSARRDRLPPRRRLLRLRRIAADPVHHLAELADRRSAIVPVHPSRAAAFLRSGRLLC
                     RRRRPGRCFIEFHGEIRESCGRSRREFITKMSPIEITPIYSLLDLDMHKMDRKVEEEE
                     EAEPMIQEMDELGLGESITDEEYLACLGNLPHGPPRPDTKTKLQGEQLHDLYIRHTVY
                     RLKYYKKSIS"
     LTR             7085. .10860
                     /note="5' LTR"
     gene            12336. .17054
                     /gene="P0458E05.2"
     CDS             12336. .17054
                     /gene="P0458E05.2"
                     /codon_start=1
                     /product="putative polyprotein"
                     /protein_id="BAC05603.1"
                     /db_xref="GI:21902055"
                     /translation="MASRPRRAARAPARFGFEDGRVEPEENHESANEQSHHSNRSHRT
                     VSRNAEVEQPRRSSRSHRTVSRDAEMGQSHRSHRTASHNSGEERLPFPPHVETNMQHL
                     MAVQTQILQGLTTAIASFQHNAHGNGHPHMGNNRSKLTDFLRSRPPEFSQTVEPVEAD
                     DWLKDVDRKLNLVQCTPVEKMLYASHQLRGPAADWWENYCNAHPEPTNIAWDEFATAF
                     RAAHVPESTIDMKKEEFNRLKQGNSSINEYLSQFNKLARYAPEEVDTDKKKIRKFLKG
                     IAVGMRLQLLAHDFPTFQHMINKALLLEDARKEATEEYKKRKSNHQGNSFRGAPRPRY
                     GQPMQYHQSVTQANRQPGYAPRPQINRPAPQPQQRAPSGNTAPNSVISFKSPQGPSAV
                     QCFRCNQMGHYARQCPQNPTNTNSGHANGSTARTPTPAAAQSRPSSQASGQGSRASNN
                     FGRGRVNHIQAETAQDAPDVVMGMFSVNSVPAIVLFDSGASHSFISQAFVKRNGWKTQ
                     NLRVPMIVHSPGRNIRATQICPEVNLRIEEVDFLAKPIVLDSQSLDIILGMDWLAKHK
                     GQIDCAEKSITLQGPGGKQVRFTPNTPTASRSILTCLQVTSLESVPIVCEYPDVFPEE
                     LKRMPPDREIEFAIELAPGTAPIAKRPYRMAANELAEVKRQIEELESKGYVRPSSSPW
                     GAPVLLVKKKDGSERMVIDYRALNEVTIKNKYPLPRIDDLFDQLKGARVFSKIDLRSG
                     YHQLKIRSEDIPKTAFSTRYGLYEFTVMSFGLTNAPAFFMNLMNKIFMEYLDQFVVVF
                     IDNILIYSKNEEEHAEHLRRIMEKLRDHQLFAKFSKCEFWLDRVAFLGHVISSNGVEV
                     DPSKVEAVLAWNPPKNVSEIRSFLGLAGYYRRFIEGFSKLARPMTELLKKEKKFQWSA
                     ACEDSFQELKKRLTTAPVLTLPDIRKDFEIFCDASRQGLGCVLMQEQKVVAYASRQLR
                     PHEANYPTHDLELAAVVHALKIWRHYLIGNRCEVYTDHKSLKYIFTQTELNMRQRKWL
                     ELIKDYDLGIHYHPGKANVVADALSHKAYCNVAQIRPDQDHLCRELEKLRLTVVQSGV
                     PASLTVQPTLESQIREAQKDDEGIKELIKRIQEKKDTNFSIDDKGTVWCGPRICVPAK
                     KELRDLILKEAHESAYSIHPGSTKMYQDIKAYFWWAGMKRDVAEYVALCDICQRVKAE
                     HQRPAGLLQPLPIPEWKWEEIGMDFITGLPRTPSGYDSIWVIVDRLTKSAHFVPVKTT
                     YDGKKLAELYMTHVVCRFGCPKKIVSDRGTQFTSRFWKQLHEALGTDLNFSTAYHPQT
                     DGQTERVNQILEDMLRACALDFEGTWDRCLPYAEFSYNNSYQASIQMSPNEAMFGRKC
                     RTPLCWNEVGEALVFGPDILKSAEEQVKLIRERLKTAQNRQKNYADNRRRDLEFEKGD
                     HVYLRVSPLRGMRRFGMSGKLAPRYIGPYLITARHGEVAYQLELPEGLADVHNVFHVS
                     QLKKCLRVPEEQVPLGNIELEKNLTYKERPIKVLEEVERQTRRKTIKFYKVQWSNHSE
                     DEATWEREDLLRAEFPDLLP"
     LTR             17098. .20885
                     /note="3' LTR"
     LTR             20886. .24214
                     /note="5' LTR"
     gene            25856. .30421
                     /gene="P0458E05.3"
     CDS             25856. .30421
                     /gene="P0458E05.3"
                     /codon_start=1
                     /product="putative polyprotein"
                     /protein_id="BAC05604.1"
                     /db_xref="GI:21902056"
                     /translation="MVYTRTGSRATGEGSNGEERADGVHPNGDSGNGPPPLPENPTLA
                     QVMAHQTQMMAAMMQQMQQQHQQMHQRMMQHAEQQHQQFGPPPPQSKLPEFLRVRPPT
                     FSSTTNPMEANDWLHAIEKKLNLLQCNDQEKVAFATHQLQGPASAWWDNHMATRPPGT
                     EVTWAEFCRSFRKAQVPDGVVAQKKREFRALHQGNRTVTEYLHEFNRLARYAPEDVRT
                     DAEKQEKFMAGLDDELTNQLISGDYADFERLVDKAIRQEDQRNKMDRKRKAAQFRSNQ
```

```
                     GSQHRPRFTPGQQGGPTTMIVRQHRPFNPSNFHQGTSGSQNHHGGQPNRGAAPRPPMA
                     PAQSGQPAQAKKETGAKPGSCFNCGELGHFADKCPKPRRAGPRFIQARVNHASAEEAQ
                     AAPEVVLGTFPVNSIPATVLFDSGATHSFISKKFVGMHGLIREELSTPMRVHTPGNSS
                     TSVQFSPSITIEIQRSPFLANLILLESKDLDVILGMDWLTKFKGVIDCASRTVTLTNE
                     KGETVVYKSLVSPKKGASLNQIETEIPVDTEEKNLRKLEDIPIVCEYPEVFPEDLTTM
                     PPKREIEFRIDLAPGTAPIYKRPYRMAANELAEVKKQVDEQLQKGYIRPSTSPWGAPV
                     IFVEKKDKTKRMCVDYRALNEVTIKNKYPLPRIDDLFDQLKGAKVFSKIDLRSGYHQL
                     RIREEDIPKTAFTTRYGLYECTVMSFGLTNAPAFFMNLMNKVFMEFLDKFVVVFIDDI
                     LIYSKSEEEHEQHLRLVLEKLKEHQLYAKFSKCDFWLTEVKFLGHVITAQGVAVDPSN
                     VESVTKWTPPKTVSQIRSFLGLAGYYRRFIENFSRIARPMTQLLKKDEKFKWTAECDK
                     SFEELKRKLVSAPVLILPDQTKDFQVYCDASRHGLGCVLMQEGRVVAYASRQLRPHEG
                     NYPTHDLELAAVVHALKIWRHYLIGNRCEVYTDHKSLKYIFTQPDLNLRQRRWLELIK
                     DYDMSIHYHPGKANVVADALSRKSYCTALCIEGMCEELRQEFEHLNVGIVEHGFVAAL
                     EARPTLVDQVRAAQVNDPEIAELKKNMRVGKARDFLEDEHGTIWMGERLCVPDDKELK
                     DLILTEAHQTPYSIHPGSTKMYQDLKEKFWWVSMRREIAEFVALCNVCQRVKAEHQRP
                     AGLLQPLQIPEWKWEEIGMDFITGLPRTSSGHDSIWVVVDRLTKVAHFIPVHTTYTGK
                     RLAELYLARIMCLHGVPKKIVSDRGSQFTSKFWQKLQEELGTRLNFSTAYHPQTDGQT
                     ERVNQILEDMLRACALDFGGAWDKSLPYAEFSYNNSYQASLQMAPFEALYGRKCRTPL
                     FWDQTGERQLFGTEVLAEAEEKVRIIRERLRIAQSRQKSYADNRRRELTFEAGDYVYL
                     RVTPLRGVHRFQTKGKLAPRFVGPYKILERRGEVAYQLELPSNMIGIHDVFHVSQLKK
                     CLRVPEEQADSEHIDIQEDLTYVEKPVRILDTSERRTRNKVTRFCRVQWSHHSEEEAT
                     WEREDELKAAHPHLFTSSSESRGRDSV"
LTR                  30434. .33762
                     /note="3' LTR"
gene                 complement(join(36988. .37110,40188. .40622))
                     /gene="P0458E05.4"
CDS                  complement(join(36988. .37110,40188. .40622))
                     /gene="P0458E05.4"
                     /note="hypothetical protein"
                     /codon_start=1
                     /protein_id="BAC05605.1"
                     /db_xref="GI:21902057"
                     /translation="MDAAGTTKISDFGLAKLLQPDQTRTFTSVRGTRGYLASEWYRGA
                     GPVMVKADVYSYGVVLLETVACRRSMEMEEAAGEEERTLAEWAYEWLLVKSEAKSAMS
                     SDETVEAAEVERVVRVAMWCVQAEPQSRPSMNGVILMLQGRLERQHHEEEAVAFGRYL
                     CGEHVLHREDDVGEFSEWQSHVLVI"
LTR                  complement(40716. .41015)
                     /note="3' LTR"
gene                 complement(41118. .45086)
                     /gene="P0458E05.5"
misc_feature         complement(41118. .45086)
                     /gene="P0458E05.5"
                     /note="probably inactive due to stop codon(s) in CDS
                     pseudogene, putative retroelement pol polyprotein"
                     /pseudo
LTR                  complement(45135. .45440)
                     /note="5' LTR"
gene                 join(56253. .58166,58250. .58411,58862. .59407,59787. .59867,
                     60240. .60320)
                     /gene="P0458E05.6"
CDS                  join(56253. .58166,58250. .58411,58862. .59407,59787. .59867,
                     60240. .60320)
                     /gene="P0458E05.6"
                     /codon_start=1
                     /product="putative chromosome condensation factor"
                     /protein_id="BAC05606.1"
                     /db_xref="GI:21902058"
```

Query Match 45.1%; Score 909.2; DB 8; Length 193577;
Best Local Similarity 79.9%; Pred. No. 1.9e-114;
Matches 1096; Conservative 1; Mismatches 269; Indels 6; Gaps 2;

```
Qy       219 ATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAG 278
Db    151179 ATATGCGGGGCGTCGCTGTTCATCAAGGAGGACATGAAGATCACGGACGTCGAGGTGGAG 151120

Qy       279 GTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACG 338
Db    151119 ATCCTGCTCGGCATCCTCAACCTCTACTCCCTCGTCGGCTCCTTCGCCGCCGGCCGGACC 151060

Qy       339 TCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGG 398
Db    151059 TCCGACTGGATCGGCCGCCGCCTCACCATCATCCTCGCCGCGGTCATCTTCTTCGTCGGC 151000

Qy       399 GSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCC 458
Db    150999 GCCATCATGATGGGCCTCTCCGTCAACTACCCCATGCTCATGGCCGGCCGCTTCGTCGCC 150940

Qy       459 GGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCG 518
Db    150939 GGCATCGGCGTCGGCTACGCCTTCATGATCGCGCCCGTCTACACCGCCGAGGTCTCGCCG 150880

Qy       519 GCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTG 578
Db    150879 GCGTCGTCGCGCGGCTTCCTCACCTCCTTCCCGGAGGTGTTCATCAACTTCGGCATCCTC 150820

Qy       579 CTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATC 638
Db    150819 CTCGGCTACGTCTCCAACTACGCCTTCTCGCGCCTCCGCCTGCAGCTCGGCTGGCGCCTC 150760

Qy       639 ATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCG 698
Db    150759 ATGCTCGGCGTCGGCGCGGCGCCGTCCGTCGCCCTCGCGCTCATGGTGCTCGCCATGCCG 150700

Qy       699 GAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAG 758
Db    150699 GAGTCGCCGCGGTGGCTCGTCATGAAGGGCCGCCTCGCCGACGCCAAGGTCGTCCTCGGC 150640

Qy       759 AAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCC 818
Db    150639 GAGACCTCGGACACGGCGGAGGAGGCCGCCACCCGCCTCGCGGAAATCAAGGAGGCGGTG 150580

Qy       819 GGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAAC 878
Db    150579 GCCATCCCGGCCGACCTCGACGGTGACGTCGTGGCCGTGCCGAAGAGGGCCGGCGG---C 150523

Qy       879 GAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTG 938
Db    150522 GAGAGGCGCGTGTGGAAGGAGCTGATCCTGTCGCCGACGCCGGCCGTCCGGCGCGTGCTC 150463

Qy       939 CTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTC 998
Db    150462 CTGTCGGCGCTCGGCATCCACTTCTTCCAGCAGTCGTCGGGCATCGACGCCGTCGTGCTC 150403

Qy       999 TACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGCACC 1058
Db    150402 TACAGCCCGCGCGTGTTCCAGAGCGCCGGGATCACCGACAAGAACAAGCTGCTCGGCACG 150343

Qy      1059 ACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGAC 1118
Db    150342 ACGTGCGCCGTGGGCGTCACCAAGACGCTCTTCATCCTGGTGGCAACGTTCACGCTCGAC 150283

Qy      1119 GGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATC 1178
Db    150282 CGCTTCGGCCGGCGGCCGCTGCTGCTGGCCAGCGCCGGCGGCATGATCGCCACCCTCGTC 150223

Qy      1179 GGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCC 1238
Db    150222 ACCCTCGGCTTAGGCCTCACCGTCATCGGTGAGGACGCCACCGGC---GGCGGCTGGGCC 150166

Qy      1239 ATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGC 1298
Db    150165 ATCGCCGTGTCCATCGCCTCCATCCTCGCCTTCGTCGCCTTCTTCTCCATCGGCCTCGGC 150106

Qy      1299 CCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGC 1358
Db    150105 CCCATCACCTGGGTGTACAGCTCCGAGATCTTCCCGCTCCACCTCCGCGCCCTCGGCTGC 150046

Qy      1359 TCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCG 1418
Db    150045 GCTCTCGGCGTCGGCCTCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTCTCC 149986

Qy      1419 CTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTC 1478
Db    149985 CTCTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACGCCGGCGTCGCCTCGCTT 149926
```

```
Qy       1479 GCCTGGGTGTTCTTCTACACCTACCTCCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATG 1538
Db     149925 GCTTGGCTCTTCTTCTTCACCTACCTCCCGGAGACTCGTGGACGCACGCTCGAGCAGATG 149866

Qy       1539 AGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAGCCA 1590
Db     149865 GGCGAGCTCTTCCGCATCCACAACATGGCCGGCGACGACGACTCTGCCGCCA 149814
```

RESULT 14
AC116426/c
LOCUS AC116426 155037 bp DNA linear PLN 28-MAR-2002
DEFINITION Genomic sequence for Oryza sativa, Nipponbare strain, clone OSJNBb0058P18, from chromosome 3, complete sequence.
ACCESSION AC116426
VERSION AC116426.1 GI:19774446
KEYWORDS HTG.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 155037)
AUTHORS McCombie,W.R., Spiegel,L., de la Bastide,M., Preston,R., Kuit,K., Ferraro,K., Nascimento,L., Zutavern,T., Balija,V., Bell,M., Baker,J., Santos,L., Miller,B., Katzenberger,F., Muller,S., King,L., Yang,C., Dike,S., O'Shaughnessy,A., Palmer,L. and Dedhia,N.
TITLE Genomic sequence for Oryza sativa, Nipponbare strain, clone OSJNBb0058P18, from chromosome 3, complete sequence
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 155037)
AUTHORS McCombie,W.R.
TITLE Direct Submission
JOURNAL Submitted (28-MAR-2002) Lita Annenberg Hazen Genome Center, Cold Spring Harbor Laboratories, 1, Bungtown Road, Cold Spring Harbor, NY 11724, USA
COMMENT This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.
Clone OSJNBb0058P18 overlaps finished clone OSJNBa0091J11 (AC115687) from base 1 to base 94350. The overlap is from base 86812 to base 181163 on OSJNBa0091J11.
FEATURES Location/Qualifiers
source 1. .155037
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="3"
/clone="OSJNBb0058P18"
/clone_lib="EcoRI"
misc_feature 3770. .3840
/note="We believe the assembly to be correct. The sequence is covered by a PCR product which was amplified with a high fidelity polymerase. There is additional coverage by a single subclone with quality below phred30."
misc_feature 16142. .16264
/note="We believe the assembly to be correct. The sequence is a polynucleotide (ATnACATn) repeat in which the exact number of copies is unknown. Two subclones span the repeat into unique sequence on both sides."
misc_feature 65646. .65667
/note="We believe the assembly to be correct. The sequence is a mononucleotide (C) repeat in which the exact number of Cs is unknown. One subclone in the region spans into unique sequence on both sides."
misc_feature 136662. .136740
/note="We believe the assembly to be correct. The sequence is covered solely by direct dye terminator sequencing reactions done with BAC template DNA. The sequence is high quality."
ORIGIN

Query Match 37.6%; Score 759.4; DB 8; Length 155037;
Best Local Similarity 73.0%; Pred. No. 4.8e-94;
Matches 992; Conservative 1; Mismatches 357; Indels 9; Gaps 1;

```
Qy        207 GATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGAC 266
Db      89013 GACATCTCGGTGATGAGCGGCGCGCAGAAGTTCATGAAGAAGGACCTGAACATCTCGGAC 88954

Qy        267 GGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCG 326
Db      88953 GCGAAGGTGGAGGTCCTCGCCGGCATCATCAACATCTACTCGCTCGTCGGCTCGCTCGCG 88894

Qy        327 GCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATA 386
Db      88893 GCGGGCCGGACGTCCGACTGGATCGGCCGTCGCTACACCATGGTGCTCGCGTCGGTCATA 88834

Qy        387 TTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGC 446
Db      88833 TTCTTCGCCGGCGCGCTCATCATGGGCCTCGCCCCGAGCTACGCGATCGTCATGCTCGGC 88774

Qy        447 CGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCC 506
Db      88773 CGCTTCGTGGCCGGCGTGGGTGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACGGCC 88714

Qy        507 GAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAAC 566
Db      88713 GAGGTGGCGCCGACGTCCGCCCGCGGCTTGCTCACGTCGTTCCCCGAGGTGTTCATCAAC 88654

Qy        567 TTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTC 626
Db      88653 GCCGGCGTCCTCCTCGGCTACGTCTCCAACTTCGCCTTCTACCGCCTCCCGCTGCACATT 88594

Qy        627 GGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTG 686
Db      88593 GGCTGGCGCGTCATGTTCCTAGTCGGCGCCGTGCCGCCCGCCTTCCTCGCCGTGGGCGTC 88534

Qy        687 CTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAG 746
Db      88533 CTCGCCATGCCGGAGTCGCCGCGGTGGCTCGTCATGCAGGGGCGCATCGGCGACGCGCGA 88474

Qy        747 GTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATC 806
Db      88473 CGTGTCCTCGAGAAGACCTCCGACTCGCCGGCCGAGGCCGAGGAGCGGCTCGCGGACATC 88414

Qy        807 AAGGCCGCCGCCGGCATCCCTGAGGAGCTC---------GACGGCGACGTGGTGACCGTC 857
Db      88413 AAGAACGCGGTCGGCATCCCCGAAGGGATCTCCGACGAGGACGAAGTTGTAGCCGTCGTC 88354

Qy        858 CCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACC 917
Db      88353 CACAAGAGCAGGGGCTCCCACGGCGAGGGTGTCTGGAGGGACCTGCTGCTCCGCCCGACG 88294

Qy        918 CCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTG 977
Db      88293 CCGGCCGTGCGCCGCATACTCATCGCCTGCCTCGGCCTCCAGTTCTTCCAGCAAGCCTCC 88234

Qy        978 GGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAAC 1037
Db      88233 GGCATCGACGCCGTCGTGCTGTACAGCCCGCGGGTGTTCGACAACGCGGGCCTCCACTCC 88174

Qy       1038 GACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTG 1097
Db      88173 GACTCCGACTCCATCGGAGCCTCCGTGGCGGTGGGCGCCAGCAAGACGCTCTTCATCCTC 88114

Qy       1098 TTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGC 1157
Db      88113 GTGGCCACGTTCCTCCTCGACCGCGTCGGCCGGAGGCCGCTGCTCCTCACCAGCGCCGGC 88054
```

```
Qy       1158 GGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCC 1217
Db      88053 GGGATGGTGATCTCGCTCGTCACGCTGGCCTCGGCGCTGCACATGATCGAGCACCGCCCG 87994

Qy       1218 GACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCC 1277
Db      87993 GAGGGCCAGGCGACGGCGCTGGTGGGTCTGAGCATCGCGATGGTGCTGGTGTTCGTGGCG 87934

Qy       1278 TTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTC 1337
Db      87933 TCCTTCTCCATCGGGATGGGCCCGATCGCGTGGGTGTACAGCTCGGAGATCTTCCCGCTG 87874

Qy       1338 CAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTC 1397
Db      87873 CGGCTGCGCGCGCAGGGATGCGCGCTCGGCACGGCGATGAACCGGGTCGTGAGCGGCGCC 87814

Qy       1398 ATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTC 1457
Db      87813 GTCAGCATGTCCTTCATCTCGCTCTACAAGGCCATCACCTTCGCCGGGAGCTTCTACCTC 87754

Qy       1458 TACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGC 1517
Db      87753 TACGCCGGCATAGCTGCGGCCGGGTGGGTGTTCATGTTCTTCTTCCTGCCGGAGACGCAA 87694

Qy       1518 GGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGAC 1556
Db      87693 GGCAGGAGCCTGGAGGACACCGTCAAACTCTTCGGCGGC 87655
```

RESULT 15
AC115687/c
LOCUS AC115687 181163 bp DNA linear PLN 22-MAR-2002
DEFINITION Genomic sequence for Oryza sativa, Nipponbare strain, clone OSJNBa0091J11, from chromosome 3, complete sequence.
ACCESSION AC115687
VERSION AC115687.1 GI:19683096
KEYWORDS HTG.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 181163)
AUTHORS McCombie,W.R., de la Bastide,M., Spiegel,L., Preston,R., Ferraro,K., Kuit,K., Nascimento,L., Zutavern,T., Balija,V., Bell,M., Baker,J., Santos,L., Miller,B., Katzenberger,F., Muller,S., King,L., Yang,C., Dike,S., O'Shaughnessy,A., Palmer,L. and Dedhia,N.
TITLE Genomic sequence for Oryza sativa, Nipponbare strain, clone OSJNBa0091J11, from chromosome 3, complete sequence
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 181163)
AUTHORS McCombie,W.R.
TITLE Direct Submission
JOURNAL Submitted (22-MAR-2002) Lita Annenberg Hazen Genome Center, Cold Spring Harbor Laboratories, 1, Bungtown Road, Cold Spring Harbor, NY 11724, USA
COMMENT This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.
FEATURES Location/Qualifiers
source 1. .181163
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="3"
/clone="OSJNBa0091J11"
/clone_lib="HindIII"
misc_feature 52283. .52374
/note="We believe the assembly to be correct. The sequence is a dinucleotide (TA) repeat followed by a dinucleotide (CA) repeat. Four subclones span into unique sequence on both sides."
misc_feature 53504. .53580
/note="We believe the assembly to be correct. The sequence is a dinucleotide (TA) repeat followed by a dinucleotide (CA) repeat. Three subclones span into unique sequence on both sides."
misc_feature 132718. .132738
/note="We believe the assembly to be correct. The sequence is a mononucleotide (C) repeat with a single unique base (A) followed by another mononucleotide (C) repeat. A single subclone spans the region into unique sequence on both sides."
misc_feature 152459. .152479
/note="We believe the assembly to be correct. The sequence is a mononucleotide (C) repeat flanking a single unique base (A) which is followed by an additional mononucleotide (C) repeat. A single subclone spans through to unique sequence on both sides."
ORIGIN

Query Match 37.6%; Score 759.4; DB 8; Length 181163;
Best Local Similarity 73.0%; Pred. No. 4.8e-94;
Matches 992; Conservative 1; Mismatches 357; Indels 9; Gaps 1;

```
Qy        207 GATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGAC 266
Db     175826 GACATCTCGGTGATGAGCGGCGCGCAGAAGTTCATGAAGAAGGACCTGAACATCTCGGAC 175767

Qy        267 GGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCG 326
Db     175766 GCGAAGGTGGAGGTCCTCGCCGGCATCATCAACATCTACTCGCTCGTCGGCTCGCTCGCG 175707

Qy        327 GCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATA 386
Db     175706 GCGGGCCGGACGTCCGACTGGATCGGCCGTCGCTACACCATGGTGCTCGCGTCGGTCATA 175647

Qy        387 TTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGC 446
Db     175646 TTCTTCGCCGGCGCGCTCATCATGGGCCTCGCCCCGAGCTACGCGATCGTCATGCTCGGC 175587

Qy        447 CGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCC 506
Db     175586 CGCTTCGTGGCCGGCGTGGGTGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACGGCC 175527

Qy        507 GAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAAC 566
Db     175526 GAGGTGGCGCCGACGTCCGCCCGCGGCTTGCTCACGTCGTTCCCCGAGGTGTTCATCAAC 175467

Qy        567 TTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTC 626
Db     175466 GCCGGCGTCCTCCTCGGCTACGTCTCCAACTTCGCCTTCTACCGCCTCCCGCTGCACATT 175407

Qy        627 GGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTG 686
Db     175406 GGCTGGCGCGTCATGTTCCTAGTCGGCGCCGTGCCGCCCGCCTTCCTCGCCGTGGGCGTC 175347

Qy        687 CTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAG 746
Db     175346 CTCGCCATGCCGGAGTCGCCGCGGTGGCTCGTCATGCAGGGGCGCATCGGCGACGCGCGA 175287

Qy        747 GTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATC 806
Db     175286 CGTGTCCTCGAGAAGACCTCCGACTCGCCGGCCGAGGCCGAGGAGCGGCTCGCGGACATC 175227

Qy        807 AAGGCCGCCGCCGGCATCCCTGAGGAGCTC---------GACGGCGACGTGGTGACCGTC 857
Db     175226 AAGAACGCGGTCGGCATCCCCGAAGGGATCTCCGACGAGGACGAAGTTGTAGCCGTCGTC 175167

Qy        858 CCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACC 917
```

```
Db    175166 CACAAGAGCAGGGGCTCCCACGGCGAGGGTGTCTGGAGGGACCTGCTGCTCCGCCCGACG 175107

Qy       918 CCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTG 977

Db    175106 CCGGCCGTGCGCCGCATACTCATCGCCTGCCTCGGCCTCCAGTTCTTCCAGCAAGCCTCC 175047

Qy       978 GGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAAC 1037

Db    175046 GGCATCGACGCCGTCGTGCTGTACAGCCCGCGGGTGTTCGACAACGCGGGCCTCCACTCC 174987

Qy      1038 GACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTG 1097

Db    174986 GACTCCGACTCCATCGGAGCCTCCGTGGCGGTGGGCGCCAGCAAGACGCTCTTCATCCTC 174927

Qy      1098 TTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGC 1157

Db    174926 GTGGCCACGTTCCTCCTCGACCGCGTCGGCCGGAGGCCGCTGCTCCTCACCAGCGCCGGC 174867

Qy      1158 GGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCC 1217

Db    174866 GGGATGGTGATCTCGCTCGTCACGCTGGCCTCGGCGCTGCACATGATCGAGCACCGCCCG 174807

Qy      1218 GACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCC 1277

Db    174806 GAGGGCCAGGCGACGGCGCTGGTGGGTCTGAGCATCGCGATGGTGCTGGTGTTCGTGGCG 174747

Qy      1278 TTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTC 1337

Db    174746 TCCTTCTCCATCGGGATGGGCCCGATCGCGTGGGTGTACAGCTCGGAGATCTTCCCGCTG 174687

Qy      1338 CAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTC 1397

Db    174686 CGGCTGCGCGCGCAGGGATGCGCGCTCGGCACGGCGATGAACCGGGTCGTGAGCGGCGCC 174627

Qy      1398 ATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTC 1457

Db    174626 GTCAGCATGTCCTTCATCTCGCTCTACAAGGCCATCACCTTCGCCGGGAGCTTCTACCTC 174567

Qy      1458 TACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGC 1517

Db    174566 TACGCCGGCATAGCTGCGGCCGGGTGGGTGTTCATGTTCTTCTTCCTGCCGGAGACGCAA 174507

Qy      1518 GGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGAC 1556

Db    174506 GGCAGGAGCCTGGAGGACACCGTCAAACTCTTCGGCGGC 174468
```

Search completed: July 2, 2004, 06:17:20
Job time : 12609 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: July 2, 2004, 00:48:19 ; Search time 1156 Seconds
(without alignments)
7412.297 Million cell updates/sec

Title: US-10-051-902A-21
Perfect score: 2017
Sequence: 1 cttacatgtaagctcgtgcc.....aaaaaaaaaaaaaaaaaaa 2017

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	2016.6	100.0	2017	6	ABK51972	Abk51972 Rice cont
	2	2016.6	100.0	2017	7	ABX93208	Abx93208 cDNA enco
	3	1094.6	54.3	2089	6	ABK51974	Abk51974 Wheat cDN
	4	1094.6	54.3	2089	7	ABX93210	Abx93210 cDNA enco
	5	1005	49.8	1914	6	ABK51971	Abk51971 Corn cDNA
	6	1005	49.8	1914	7	ABX93207	Abx93207 cDNA enco
	7	981.6	48.7	1872	6	ABK51975	Abk51975 Wheat cDN
	8	981.6	48.7	1872	7	ABX93211	Abx93211 cDNA enco
	9	575.2	28.5	1853	6	ABK51973	Abk51973 Soybean c
	10	575.2	28.5	1853	7	ABX93209	Abx93209 cDNA enco
	11	477.2	23.7	2056	6	ABL41873	Abl41873 Nucleotid
	12	470.2	23.3	1527	3	AAC43261	Aac43261 Arabidops
	13	469.2	23.3	1925	6	ABL41874	Abl41874 Nucleotid
c	14	443.2	22.0	1662	6	ABL41878	Abl41878 Nucleotid
	15	433.8	21.5	1482	6	ABZ12990	Abz12990 Arabidops
	16	426	21.1	1766	6	ABL41872	Abl41872 Nucleotid
c	17	397.4	19.7	1715	6	ABL41875	Abl41875 Nucleotid
	18	387.6	19.2	1722	6	ABL41876	Abl41876 Nucleotid
	19	343.6	17.0	615	6	ABK51970	Abk51970 Corn cDNA
	20	343.6	17.0	615	7	ABX93206	Abx93206 cDNA enco
	21	325.4	16.1	1644	6	ABZ12515	Abz12515 Arabidops
c	22	325	16.1	1690	6	ABL41879	Abl41879 Nucleotid
	23	313.2	15.5	756	8	ADA49209	Ada49209 Maize gen
	24	295	14.6	1004	8	ADA48319	Ada48319 Rice gene
c	25	191.2	9.5	2000	7	ADA72262	Ada72262 Rice gene
	26	188.8	9.4	1675	2	AAZ32195	Aaz32195 Corn hexo
	27	172.4	8.5	1560	7	ADA70449	Ada70449 Rice gene
	28	170	8.4	1377	7	ADA70153	Ada70153 Rice gene
	29	168.4	8.3	1776	2	AAZ32200	Aaz32200 Rice hexo
	30	164.2	8.1	616	8	ADA49126	Ada49126 Wheat gen
c	31	159.8	7.9	349980	6	ABQ81849	Abq81849 Bifidobac
	32	158.2	7.8	1539	7	ADA70241	Ada70241 Rice gene
	33	152.8	7.6	1350	7	ACA37723	Aca37723 Prokaryot
c	34	148.8	7.4	1545	7	ACA40941	Aca40941 Prokaryot
	35	147.2	7.3	1509	7	ACA40810	Aca40810 Prokaryot
	36	147.2	7.3	110000	4	AAI99682_37	Continuation (38 o
	37	141.2	7.0	1752	2	AAZ32203	Aaz32203 Wheat hex
	38	139.8	6.9	1572	7	ADA70338	Ada70338 Rice gene
	39	136.4	6.8	1577	8	ADA48271	Ada48271 Rice gene
	40	136.4	6.8	1577	9	ADC07783	Adc07783 Rice DNA
	41	136	6.7	1167	2	AAZ32199	Aaz32199 Rice hexo
	42	130.2	6.5	1395	7	ACA32312	Aca32312 Prokaryot
	43	129.8	6.4	1281	7	ACA38748	Aca38748 Prokaryot
	44	128	6.3	110000	4	AAI99683_37	Continuation (38 o
	45	121.4	6.0	2504	4	AAF55869	Aaf55869 Rat GLUTX

ALIGNMENTS

RESULT 1
ABK51972
ID ABK51972 standard; cDNA; 2017 BP.
XX
AC ABK51972;
XX
DT 27-AUG-2002 (first entry)
XX
DE Rice contig encoding Beta vulgaris-like sugar transport protein.
XX
KW Rice; Beta vulgaris-like sugar transport protein; carbohydrate transport;
KW grain filling; annual field crop; plant; gene; ss.
XX
OS Oryza sativa.
XX
FH Key Location/Qualifiers
FT CDS 96. .1628
FT /*tag= a
FT /product= "Rice Beta vulgaris-like sugar transport
FT protein"
XX
PN US6383776-B1.
XX
PD 07-MAY-2002.
XX
PF 14-APR-1999; 99US-00291922.
XX
PR 24-APR-1998; 98US-0083044P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX
DR WPI; 2002-453364/48.
DR P-PSDB; AAU97211.
XX
PT New nucleic acid encoding plant sugar-transport proteins, useful for
PT preparing transgenic plants with altered carbohydrate distribution.
XX
PS Example 4; Col 61-64; 54pp; English.
XX
CC The present invention relates to the isolation of plant polynucleotide
CC sequences encoding an Arabidopsis thaliana-like sugar transport protein
CC or Beta vulgaris-like sugar transport protein. The polynucleotide
CC sequences are useful for altering the level of sugar transport proteins
CC in plants, i.e. for control of carbohydrate transport and distribution in

CC plant cells, e.g. during grain filling of annual field crops (e.g. corn, rice, soybeans, and wheat), and, for studying carbohydrate flows and sugar transport. The polynucleotide sequences can also be used to isolate cDNA sequences and genes that encode homologues of the new proteins. The present sequence representing a contig assembled from various rice cDNA clones encodes a Beta vulgaris-like sugar transport protein
XX
SQ Sequence 2017 BP; 389 A; 627 C; 582 G; 418 T; 0 U; 1 Other;

Query Match 100.0%; Score 2016.6; DB 6; Length 2017;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2017; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CTTACATGTAAGCTCGTGCCGGCACGAGCTTACACTCGACCGCCACTACTGTACACGGCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTTACATGTAAGCTCGTGCCGGCACGAGCTTACACTCGACCGCCACTACTGTACACGGCC 60

Qy         61 CAGAGCGAGCCTCCTCCTCCTCTGCACCACCGGAGATGGCTTCCGCCGCGCTGCCGGAGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGAGCGAGCCTCCTCCTCCTCTGCACCACCGGAGATGGCTTCCGCCGCGCTGCCGGAGG 120

Qy        121 CCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCT 180

Qy        181 CCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACA 240

Qy        241 TCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACC 300

Qy        301 TCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGT 360

Qy        361 ACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCG 420

Qy        421 TCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGC 480

Qy        481 TCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGA 540

Qy        541 CGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATG 600

Qy        601 CTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGC 660

Qy        661 CGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCA 720

Qy        721 TGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGG 780

Qy        781 AGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACG 840

Qy        841 GCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGC 900

Qy        901 TCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACT 960

Qy        961 TCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGA 1020

Qy       1021 GCCCCGGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GCCCCGGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCA 1080

Qy       1081 AGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGT 1140

Qy       1141 TGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCG 1200

Qy       1201 TCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCA 1260

Qy       1261 CCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCT 1320

Qy       1321 CGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACC 1380

Qy       1381 GCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 GCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCG 1440

Qy       1441 GCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCT 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCT 1500

Qy       1501 ACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 ACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGG 1560

Qy       1561 CCGCCGCCTCGGAATCAGACGAGCCAGCCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCA 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CCGCCGCCTCGGAATCAGACGAGCCAGCCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCA 1620

Qy       1621 CTAACTGATCAAACTAACCGCAAAATCACCAAATCCTAAGGGTTTTCTTGCAAAAACGTG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 CTAACTGATCAAACTAACCGCAAAATCACCAAATCCTAAGGGTTTTCTTGCAAAAACGTG 1680

Qy       1681 TGCTGTACTGGCTAGCTAGCAAGTAGTAGCAGCAACGTGGGAAGATTCGCTGATCCGGCG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 TGCTGTACTGGCTAGCTAGCAAGTAGTAGCAGCAACGTGGGAAGATTCGCTGATCCGGCG 1740

Qy       1741 TTGCTGGAGAGCGACGGCCGGCGACGACAAAGCTGAGCTCCAGCTCGAGACTTCTTAAAA 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TTGCTGGAGAGCGACGGCCGGCGACGACAAAGCTGAGCTCCAGCTCGAGACTTCTTAAAA 1800

Qy       1801 TCATCTTCAAGTACATGGATTTTATTTTGCTCTTTGCTTTGTCCGTAAAAGTTGTACTAT 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 TCATCTTCAAGTACATGGATTTTATTTTGCTCTTTGCTTTGTCCGTAAAAGTTGTACTAT 1860

Qy       1861 GCGATGAAGAATACCAGTATGTAGCAAGGCTGAGGTTGTGTGTAGCTACTAGAAGTGTCA 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 GCGATGAAGAATACCAGTATGTAGCAAGGCTGAGGTTGTGTGTAGCTACTAGAAGTGTCA 1920

Qy       1921 GTCACGTTGTTCTTGTAAGAAATGTTTAACTGTTAATTAAGCAGTATTGTTGCAGTAATC 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GTCACGTTGTTCTTGTAAGAAATGTTTAACTGTTAATTAAGCAGTATTGTTGCAGTAATC 1980

Qy       1981 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2017
```

```
Db       1981 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2017


RESULT 2
ABX93208
ID   ABX93208 standard; cDNA; 2017 BP.
XX
AC   ABX93208;
XX
DT   29-MAY-2003  (first entry)
XX
DE   cDNA encoding rice sugar transport protein #3.
XX
KW   Beta vulgaris-like sugar transport protein; corn; rice; wheat;
KW   plant sugar transport protein; carbohydrate transport; soybean;
KW   carbohydrate distribution; plant; gene; ss.
XX
OS   Oryza sativa.
XX
PN   US2002178468-A1.
XX
PD   28-NOV-2002.
XX
PF   17-JAN-2002; 2002US-00051902.
XX
PR   24-APR-1998;   98US-0083044P.
PR   14-APR-1999;   99US-00291922.
XX
PA   (ALLE/) ALLEN S M.
PA   (HITZ/) HITZ W D.
PA   (KINN/) KINNEY A J.
PA   (TING/) TINGEY S V.
XX
PI   Allen SM,  Hitz WD,  Kinney AJ,  Tingey SV;
XX
DR   WPI; 2003-340957/32.
DR   P-PSDB; ABU08336.
XX
PT   Novel plant sugar transport proteins and nucleic acid encoding the
PT   protein useful for producing transgenic plants having altered levels of
PT   sugar transport protein.
XX
PS   Claim 7; Page 32-33; 56pp; English.
XX
CC   The present invention relates to the isolation of Arabidopsis thaliana-
CC   like or Beta vulgaris-like sugar transport proteins, and the
CC   polynucleotide sequences encoding them. The plant sugar transport
CC   proteins of the invention have been isolated from corn, rice, soybean,
CC   and wheat. The polypeptides of the invention may be used for altering the
CC   level of expression of a sugar transport protein in a host cell, by
CC   transforming a host cell with a chimeric construct encoding all, or a
CC   portion of the sugar transport protein, in sense or antisense
CC   orientation. Particularly, the polypeptides may provide a means to
CC   control carbohydrate transport and distribution in plants. ABX93206-
CC   ABX93211 represent cDNA sequences encoding Beta vulgaris-like sugar
CC   transport proteins
XX
SQ   Sequence 2017 BP; 389 A; 627 C; 582 G; 418 T; 0 U; 1 Other;

  Query Match             100.0%;  Score 2016.6;  DB 7;  Length 2017;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2017;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTTACATGTAAGCTCGTGCCGGCACGAGCTTACACTCGACCGCCACTACTGTACACGGCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTTACATGTAAGCTCGTGCCGGCACGAGCTTACACTCGACCGCCACTACTGTACACGGCC 60

Qy         61 CAGAGCGAGCCTCCTCCTCCTCTGCACCACCGGAGATGGCTTCCGCCGCGCTGCCGGAGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGAGCGAGCCTCCTCCTCCTCTGCACCACCGGAGATGGCTTCCGCCGCGCTGCCGGAGG 120

Qy        121 CCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCT 180

Qy        181 CCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACA 240

Qy        241 TCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACC 300

Qy        301 TCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGT 360

Qy        361 ACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCG 420

Qy        421 TCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGC 480

Qy        481 TCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGA 540

Qy        541 CGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATG 600

Qy        601 CTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGC 660

Qy        661 CGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCA 720

Qy        721 TGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGG 780

Qy        781 AGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACG 840

Qy        841 GCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGC 900

Qy        901 TCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACT 960

Qy        961 TCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGA 1020

Qy       1021 GCCCCGGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GCCCCGGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCA 1080

Qy       1081 AGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGT 1140

Qy       1141 TGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCG 1200

Qy       1201 TCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1201 TCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCA 1260

Qy       1261 CCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCT 1320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCT 1320

Qy       1321 CGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACC 1380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACC 1380

Qy       1381 GCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCG 1440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 GCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCG 1440

Qy       1441 GCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCT 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCT 1500

Qy       1501 ACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGG 1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 ACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGG 1560

Qy       1561 CCGCCGCCTCGGAATCAGACGAGCCAGCCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCA 1620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CCGCCGCCTCGGAATCAGACGAGCCAGCCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCA 1620

Qy       1621 CTAACTGATCAAACTAACCGCAAAATCACCAAATCCTAAGGGTTTTCTTGCAAAAACGTG 1680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 CTAACTGATCAAACTAACCGCAAAATCACCAAATCCTAAGGGTTTTCTTGCAAAAACGTG 1680

Qy       1681 TGCTGTACTGGCTAGCTAGCAAGTAGTAGCAGCAACGTGGGAAGATTCGCTGATCCGGCG 1740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 TGCTGTACTGGCTAGCTAGCAAGTAGTAGCAGCAACGTGGGAAGATTCGCTGATCCGGCG 1740

Qy       1741 TTGCTGGAGAGCGACGGCCGGCGACGACAAAGCTGAGCTCCAGCTCGAGACTTCTTAAAA 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TTGCTGGAGAGCGACGGCCGGCGACGACAAAGCTGAGCTCCAGCTCGAGACTTCTTAAAA 1800

Qy       1801 TCATCTTCAAGTACATGGATTTTATTTTGCTCTTTGCTTTGTCCGTAAAAGTTGTACTAT 1860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 TCATCTTCAAGTACATGGATTTTATTTTGCTCTTTGCTTTGTCCGTAAAAGTTGTACTAT 1860

Qy       1861 GCGATGAAGAATACCAGTATGTAGCAAGGCTGAGGTTGTGTGTAGCTACTAGAAGTGTCA 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 GCGATGAAGAATACCAGTATGTAGCAAGGCTGAGGTTGTGTGTAGCTACTAGAAGTGTCA 1920

Qy       1921 GTCACGTTGTTCTTGTAAGAAATGTTTAACTGTTAATTAAGCAGTATTGTTGCAGTAATC 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GTCACGTTGTTCTTGTAAGAAATGTTTAACTGTTAATTAAGCAGTATTGTTGCAGTAATC 1980

Qy       1981 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2017
             |||||||||||||||||||||||||||||||||||||||
Db       1981 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2017


RESULT 3
ABK51974
ID   ABK51974 standard; cDNA; 2089 BP.
XX
AC   ABK51974;
XX
DT   07-AUG-2002  (first entry)
XX
DE   Wheat cDNA clone wlk8.pk0001.a11 encoding sugar transport protein.
XX
KW   Wheat; Beta vulgaris-like sugar transport protein;
KW   carbohydrate transport; grain filling; annual field crop; plant;
KW   clone wlk8.pk0001.a11; gene; ss.
XX
OS   Triticum aestivum.
XX
FH   Key             Location/Qualifiers
FT   CDS             2. .1621
FT                   /*tag=  a
FT                   /partial
FT                   /product= "Wheat Beta vulgaris-like sugar transport
FT                   protein"
FT                   /note= "This sequence lacks a start codon"
XX
PN   US6383776-B1.
XX
PD   07-MAY-2002.
XX
PF   14-APR-1999;   99US-00291922.
XX
PR   24-APR-1998;   98US-0083044P.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Allen SM,  Hitz WD,  Kinney AJ,  Tingey SV;
XX
DR   WPI; 2002-453364/48.
DR   P-PSDB; AAU97213.
XX
PT   New nucleic acid encoding plant sugar-transport proteins, useful for
PT   preparing transgenic plants with altered carbohydrate distribution.
XX
PS   Example 4; Col 71-72; 54pp; English.
XX
CC   The present invention relates to the isolation of plant polynucleotide
CC   sequences encoding an Arabidopsis thaliana-like sugar transport protein
CC   or Beta vulgaris-like sugar transport protein. The polynucleotide
CC   sequences are useful for altering the level of sugar transport proteins
CC   in plants, i.e. for control of carbohydrate transport and distribution in
CC   plant cells, e.g. during grain filling of annual field crops (e.g. corn,
CC   rice, soybeans, and wheat), and, for studying carbohydrate flows and
CC   sugar transport. The polynucleotide sequences can also be used to isolate
CC   cDNA sequences and genes that encode homologues of the new proteins. The
CC   present sequence encodes a wheat Beta vulgaris-like sugar transport
CC   protein
XX
SQ   Sequence 2089 BP; 395 A; 672 C; 605 G; 417 T; 0 U; 0 Other;

  Query Match            54.3%;  Score 1094.6;  DB 6;  Length 2089;
  Best Local Similarity  79.1%;  Pred. No. 1.9e-177;
  Matches 1343;  Conservative    1;  Mismatches  335;  Indels   18;  Gaps    3;

Qy         84 GCACCACCGGAGATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGC 143
              ||  |  ||  |||||    |||||| || |||| ||||||| ||| ||||||||||||
Db         65 GCGTCGTCGACGATGGACCGCGCCGCACTCCCGGCGGCCGTCGAGCCCAAGAAGAAGGGC 124

Qy        144 AACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGC 203
              |||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 AACGTGAGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGC 184

Qy        204 TACGATATCGGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGT 263
              ||||| ||||| |||||||||| |||||||||||||| ||||||| | || ||||
Db        185 TACGACATCGGCGTGATGAGCGGAGCGTCGCTGTACATCCAGAAGGATCTGAAGATCAAC 244

Qy        264 GACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTC 323
              |||    || ||||||| |||||||||| || ||| | |||||||||| |||||||||
Db        245 GACACCCAGCTGGAGGTCCTCATGGGCATCCTCAACGTGTACTCGCTCATTGGCTCCTTC 304

Qy        324 GCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTC 383
              |||||||||||||||| |||||||||||||||| | ||||||||| ||||||||||||
Db        305 GCGGCGGGGCGGACGTCCGACTGGATCGGCCGGCGCTTCACCATCGTCTTCGCCGCCGTC 364

Qy        384 ATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTC 443
              ||  |||||||||||  |:  || ||||||| ||| |||||||||||||||||||||||
Db        365 ATCTTCTTCGCGGGCGCCCTCATCATGGGCTTCTCCGTCAACTACGCCATGCTCATGTTC 424

Qy        444 GGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACC 503
              || |||||||||||||||||||||||||||||| |||| |||||||||||| ||| ||||
Db        425 GGGCGCTTCGTGGCCGGCATCGGCGTGGGGTACGCTCTCATGATCGCGCCCGTGAACACG 484

Qy        504 GCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATC 563
```

```
Db        485 GGCGAGGTGTCCCCCGCGTCTGCCCGTGGGGTTCTCACATCCTTCCCGGAGGTGTTCATC 544
Qy        564 AACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAAC 623
Db        545 AACTTCGGCATCCTCCTCGGATATGTCTCCAACTTCGCCTTCGCCCGCCTCTCCCTCCGC 604
Qy        624 CTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATG 683
Db        605 CTCGGCTGGCGCATTATGCTCGGCATAGGCGCGGTGCCGTCCGTCCTGCTCGCGTTCATG 664
Qy        684 GTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCC 743
Db        665 GTGCTCGGCATGCCCGAGTCTCCCCGGTGGCTCGTCATGAAGGGCCGTCTCGCGGACGCC 724
Qy        744 AAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGAC 803
Db        725 AAGGTTGTGCTTGCCAAGACGTCCGACACGCCGGAAGAGGCCGCCGAGCGCATCGCCGAC 784
Qy        804 ATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCC--- 860
Db        785 ATTAAGACTGCCGCCGGCATCCCTCTGGGCCTCGACGGCGACGTGGTCCCCGTGCCCAAA 844
Qy        861 AAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCG 920
Db        845 AACAAAGGAAGCAGCGAGGAGAAGCGCGTTTTGAAGGACCTCATCCTGTCACCGACCATA 904
Qy        921 GCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGC 980
Db        905 GCCATGCGCCACATCCTCATCGCGGGAATCGGCATCCACTTCTTCCAGCAGTCTTCGGGC 964
Qy        981 ATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGAC 1040
Db        965 ATCGACGCCGTCGTGCTCTACAGCCCGCTAGTTTTCAAGAGCGCCGGCATCACGGGCGAC 1024
Qy       1041 AAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTG 1100
Db       1025 AGCCGTCTCCGCGGCACCACCGTGGCGGTCGGGGCCACCAATACGGTCTTCATCCTGGTG 1084
Qy       1101 GCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGG 1160
Db       1085 GCCACCTTCCTCCTCGACCGCATCCGCCGGCGGCCGCTGGTGCTGACCAGCACGGGCGGC 1144
Qy       1161 ATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGAC 1220
Db       1145 ATGCTCGTCTCCTTAGTGGGCCTCGCGACGGGGCTCACCGTCATCAGCCGCCACCCGGAC 1204
Qy       1221 GCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTC 1280
Db       1205 GAGAAGATCACCTGGGCCATCGTCCTGTGCATCTTCTGCATCATGGCCTACGTGGCCTTC 1264
Qy       1281 TTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAG 1340
Db       1265 TTCTCCATCGGCCTCGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTGCAC 1324
Qy       1341 GTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATC 1400
Db       1325 GTGCGCGCGCTGGGCTGCTCCCTGGGCGTGGCCGTCAACCGCCTGACCAGCGGCGTGATC 1384
Qy       1401 TCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTAC 1460
Db       1385 TCCATGACCTTCATTTCGCTGTCCAAGGCCATGACCATCGGCGGCGCCTTCTTCCTCTTC 1444
Qy       1461 TCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGC 1520
Db       1445 GCCGGCATCGCCTCATTCGCATGGGTGTTCTTCTTCGCCTACCTGCCGGAGACCCGCGGC 1504
Qy       1521 CGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGAC 1580
Db       1505 CGCACGCTGGAGGACATGAGCTCGCTGTTCGGCAACACGGCCACGCACAAGCAGGGCGCC 1564
Qy       1581 GAGCCAGCCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCACTAACTGATCAAACTAACCG 1640
Db       1565 GCGGAAGCCGACGACGACGCCGGGGAGAAGAAGGTGGAAATGGCCGCCACCAACTGACCG 1624
Qy       1641 CAAAATCACCAAATCCTAAGGGTTTTCTTGCAAAAACGTGTGCTGTACTGGCTAGCTAGC 1700
Db       1625 CAAGTTGGCAGATCGCGATGCGAAGACTTGCGCTGTATCCGTCTCGGCTAGCTAGCTGCC 1684
Qy       1701 A-----------AGTAGTAGCAGCAACGTGGGAAGATTCGCTGATCCGGC----GTTGCT 1745
Db       1685 ACAAGGCCACATAGATGACGAAGTAGCGTGGGAAGATTCGCTGATCCGGCCGGAGCTGCC 1744
Qy       1746 GGAGAGCGACGGCCGGC 1762
Db       1745 GGAGGGCGACGGCAAGC 1761
```

RESULT 4
ABX93210
ID ABX93210 standard; cDNA; 2089 BP.
XX
AC ABX93210;
XX
DT 29-MAY-2003 (first entry)
XX
DE cDNA encoding wheat sugar transport protein #4.
XX
KW Beta vulgaris-like sugar transport protein; corn; rice; wheat;
KW plant sugar transport protein; carbohydrate transport; soybean;
KW carbohydrate distribution; plant; gene; ss.
XX
OS Triticum aestivum.
XX
PN US2002178468-A1.
XX
PD 28-NOV-2002.
XX
PF 17-JAN-2002; 2002US-00051902.
XX
PR 24-APR-1998; 98US-0083044P.
PR 14-APR-1999; 99US-00291922.
XX
PA (ALLE/) ALLEN S M.
PA (HITZ/) HITZ W D.
PA (KINN/) KINNEY A J.
PA (TING/) TINGEY S V.
XX
PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX
DR WPI; 2003-340957/32.
DR P-PSDB; ABU08338.
XX
PT Novel plant sugar transport proteins and nucleic acid encoding the
PT protein useful for producing transgenic plants having altered levels of
PT sugar transport protein.
XX
PS Claim 7; Page 37-38; 56pp; English.
XX
CC The present invention relates to the isolation of Arabidopsis thaliana-
CC like or Beta vulgaris-like sugar transport proteins, and the
CC polynucleotide sequences encoding them. The plant sugar transport
CC proteins of the invention have been isolated from corn, rice, soybean,
CC and wheat. The polypeptides of the invention may be used for altering the
CC level of expression of a sugar transport protein in a host cell, by
CC transforming a host cell with a chimeric construct encoding all, or a
CC portion of the sugar transport protein, in sense or antisense
CC orientation. Particularly, the polypeptides may provide a means to
CC control carbohydrate transport and distribution in plants. ABX93206-
CC ABX93211 represent cDNA sequences encoding Beta vulgaris-like sugar
CC transport proteins
XX
SQ Sequence 2089 BP; 395 A; 672 C; 605 G; 417 T; 0 U; 0 Other;

Query Match 54.3%; Score 1094.6; DB 7; Length 2089;
Best Local Similarity 79.1%; Pred. No. 1.9e-177;

```
  Matches 1343;  Conservative    1;  Mismatches  335;  Indels   18;  Gaps    3;

Qy         84 GCACCACCGGAGATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGC 143
Db         65 GCGTCGTCGACGATGGACCGCGCCGCACTCCCGGCGGCCGTCGAGCCCAAGAAGAAGGGC 124

Qy        144 AACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGC 203
Db        125 AACGTGAGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGC 184

Qy        204 TACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGT 263
Db        185 TACGACATCGGCGTGATGAGCGGAGCGTCGCTGTACATCCAGAAGGATCTGAAGATCAAC 244

Qy        264 GACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTC 323
Db        245 GACACCCAGCTGGAGGTCCTCATGGGCATCCTCAACGTGTACTCGCTCATTGGCTCCTTC 304

Qy        324 GCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTC 383
Db        305 GCGGCGGGGCGGACGTCCGACTGGATCGGCCGGCGCTTCACCATCGTCTTCGCCGCCGTC 364

Qy        384 ATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTC 443
Db        365 ATCTTCTTCGCGGGCGCCCTCATCATGGGCTTCTCCGTCAACTACGCCATGCTCATGTTC 424

Qy        444 GGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACC 503
Db        425 GGGCGCTTCGTGGCCGGCATCGGCGTGGGGTACGCTCTCATGATCGCGCCCGTGAACACG 484

Qy        504 GCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATC 563
Db        485 GGCGAGGTGTCCCCCGCGTCTGCCCGTGGGGTTCTCACATCCTTCCCGGAGGTGTTCATC 544

Qy        564 AACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAAC 623
Db        545 AACTTCGGCATCCTCCTCGGATATGTCTCCAACTTCGCCTTCGCCCGCCTCTCCCTCCGC 604

Qy        624 CTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATG 683
Db        605 CTCGGCTGGCGCATTATGCTCGGCATAGGCGCGGTGCCGTCCGTCCTGCTCGCGTTCATG 664

Qy        684 GTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCC 743
Db        665 GTGCTCGGCATGCCCGAGTCTCCCCGGTGGCTCGTCATGAAGGGCCGTCTCGCGGACGCC 724

Qy        744 AAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGAC 803
Db        725 AAGGTTGTGCTTGCCAAGACGTCCGACACGCCGGAAGAGGCCGCCGAGCGCATCGCCGAC 784

Qy        804 ATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCC--- 860
Db        785 ATTAAGACTGCCGCCGGCATCCCTCTGGGCCTCGACGGCGACGTGGTCCCCGTGCCCAAA 844

Qy        861 AAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCG 920
Db        845 AACAAAGGAAGCAGCGAGGAGAAGCGCGTTTTGAAGGACCTCATCCTGTCACCGACCATA 904

Qy        921 GCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGC 980
Db        905 GCCATGCGCCACATCCTCATCGCGGGAATCGGCATCCACTTCTTCCAGCAGTCTTCGGGC 964

Qy        981 ATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGAC 1040
Db        965 ATCGACGCCGTCGTGCTCTACAGCCCGCTAGTTTTCAAGAGCGCCGGCATCACGGGCGAC 1024

Qy       1041 AAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTG 1100
Db       1025 AGCCGTCTCCGCGGCACCACCGTGGCGGTCGGGGCCACCAATACGGTCTTCATCCTGGTG 1084

Qy       1101 GCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGG 1160
Db       1085 GCCACCTTCCTCCTCGACCGCATCCGCCGGCGGCCGCTGGTGCTGACCAGCACGGGCGGC 1144

Qy       1161 ATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGAC 1220
Db       1145 ATGCTCGTCTCCTTAGTGGGCCTCGCGACGGGGCTCACCGTCATCAGCCGCCACCCGGAC 1204

Qy       1221 GCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTC 1280
Db       1205 GAGAAGATCACCTGGGCCATCGTCCTGTGCATCTTCTGCATCATGGCCTACGTGGCCTTC 1264

Qy       1281 TTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAG 1340
Db       1265 TTCTCCATCGGCCTCGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTGCAC 1324

Qy       1341 GTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATC 1400
Db       1325 GTGCGCGCGCTGGGCTGCTCCCTGGGCGTGGCCGTCAACCGCCTGACCAGCGGCGTGATC 1384

Qy       1401 TCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTAC 1460
Db       1385 TCCATGACCTTCATTTCGCTGTCCAAGGCCATGACCATCGGCGGCGCCTTCTTCCTCTTC 1444

Qy       1461 TCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGC 1520
Db       1445 GCCGGCATCGCCTCATTCGCATGGGTGTTCTTCTTCGCCTACCTGCCGGAGACCCGCGGC 1504

Qy       1521 CGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGAC 1580
Db       1505 CGCACGCTGGAGGACATGAGCTCGCTGTTCGGCAACACGGCCACGCACAAGCAGGGCGCC 1564

Qy       1581 GAGCCAGCCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCACTAACTGATCAAACTAACCG 1640
Db       1565 GCGGAAGCCGACGACGACGCCGGGGAGAAGAAGGTGGAAATGGCCGCCACCAACTGACCG 1624

Qy       1641 CAAAATCACCAAATCCTAAGGGTTTTCTTGCAAAAACGTGTGCTGTACTGGCTAGCTAGC 1700
Db       1625 CAAGTTGGCAGATCGCGATGCGAAGACTTGCGCTGTATCCGTCTCGGCTAGCTAGCTGCC 1684

Qy       1701 A-----------AGTAGTAGCAGCAACGTGGGAAGATTCGCTGATCCGGC----GTTGCT 1745
Db       1685 ACAAGGCCACATAGATGACGAAGTAGCGTGGGAAGATTCGCTGATCCGGCCGGAGCTGCC 1744

Qy       1746 GGAGAGCGACGGCCGGC 1762
Db       1745 GGAGGGCGACGGCAAGC 1761
```

RESULT 5
ABK51971
ID ABK51971 standard; cDNA; 1914 BP.
XX
AC ABK51971;
XX
DT 27-AUG-2002 (first entry)
XX
DE Corn cDNA clone cepe7.pk0018.g3 encoding sugar transport protein.
XX
KW Corn; Beta vulgaris-like sugar transport protein; carbohydrate transport;
KW grain filling; annual field crop; plant; clone cepe7.pk0018.g3; gene; ss.
XX
OS Zea mays.
XX
FH Key   Location/Qualifiers
FT CDS   56. .1597
FT   /*tag= a
FT   /product= "Corn Beta vulgaris-like sugar transport
FT   protein"
XX
PN US6383776-B1.
XX
PD 07-MAY-2002.
XX
PF 14-APR-1999; 99US-00291922.
XX

PR 24-APR-1998; 98US-0083044P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX
DR WPI; 2002-453364/48.
DR P-PSDB; AAU97210.
XX
PT New nucleic acid encoding plant sugar-transport proteins, useful for
PT preparing transgenic plants with altered carbohydrate distribution.
XX
PS Example 4; Col 57-60; 54pp; English.
XX
CC The present invention relates to the isolation of plant polynucleotide
CC sequences encoding an Arabidopsis thaliana-like sugar transport protein
CC or Beta vulgaris-like sugar transport protein. The polynucleotide
CC sequences are useful for altering the level of sugar transport proteins
CC in plants, i.e. for control of carbohydrate transport and distribution in
CC plant cells, e.g. during grain filling of annual field crops (e.g. corn,
CC rice, soybeans, and wheat), and, for studying carbohydrate flows and
CC sugar transport. The polynucleotide sequences can also be used to isolate
CC cDNA sequences and genes that encode homologues of the new proteins. The
CC present sequence encodes a corn Beta vulgaris-like sugar transport
CC protein
XX
SQ Sequence 1914 BP; 343 A; 591 C; 572 G; 408 T; 0 U; 0 Other;

Query Match 49.8%; Score 1005; DB 6; Length 1914;
Best Local Similarity 79.2%; Pred. No. 3.5e-162;
Matches 1220; Conservative 1; Mismatches 311; Indels 9; Gaps 2;

```
Qy         95 GATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTT 154
              |||||||||||  || |||   |  |||||||| ||| | ||||||||||||||||  ||
Db         55 GATGGCTTCCGACGAGCTCGCAAAGGCCGTCGAGCCCAGGAAGAAGGGCAACGTCAAGTA 114

Qy        155 CGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGG 214
              ||||  |   ||  |||||||| |||||||| ||||  ||  ||||  |||||  ||  ||  ||
Db        115 TGCCTCCATATGTGCCATCCTGGCCTCCATGGCCTCTGTCATCCTTGGCTATGACATTGG 174

Qy        215 GGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGT 274
              |||||||||  ||  |||  |  |||||||||||||||||  |  ||  ||||   ||||  |  ||  |
Db        175 GGTGATGAGTGGAGCGGCCATGTACATCAAGAAGGACCTGAATATCACGGACGTGCAGCT 234

Qy        275 GGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCG 334
              ||||  |  ||  ||  ||  ||  ||  |  ||||||||||   ||||  ||||||||  |   |  |||
Db        235 GGAGATCCTGATCGGGATCCTCAGTCTCTACTCGCTGTTCGGATCCTTCGCTGGCGCGCG 294

Qy        335 GACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGC 394
              ||||||  |||  |||||||  ||  ||  |  |||  ||||||||||||||  |||||  |||||||
Db        295 GACGTCCGACAGGATCGGGCGCCGCTTGACCGTCGTGTTCGCCGCTGTCATCTTCTTCGT 354

Qy        395 GGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGT 454
              |||  :|||  ||||||||  ||||||||||||||||  |||||||||    ||||||||||
Db        355 GGGCTCGTTGCTCATGGGTTTCGCCGTCAACTACGGCATGCTCATGGCGGGCCGCTTCGT 414

Qy        455 GGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTC 514
              ||||||  ||||  ||||||||||  |  ||||||||||  ||||||||  ||||||  |  ||
Db        415 GGCCGGAGTCGGTGTGGGCTACGGGGGCATGATCGCGCCCGTGTACACGGCCGAGATCTC 474

Qy        515 GCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCAT 574
              |||  |||  ||  |  ||||||||||||||  |  ||||||||||||||||||||  ||||||||
Db        475 GCCTGCGGCGTCCCGTGGCTTCCTGACCACCTTCCCGGAGGTGTTCATCAACATCGGCAT 534

Qy        575 CCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCG 634
              ||||||  ||  |||  |  ||  ||||   ||  |||  |  |||  |  |||||   |||||||  |||||
Db        535 CCTGCTTGGCTACCTGTCCAACTTCGCGTTCGCGCGCCTCCCGCTCCACCTCGGCTGGCG 594

Qy        635 CATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCAT 694
              |  |||||||||  |||  |||||  |   |||||||   ||||||||||||  |||||  ||  ||||
Db        595 CGTCATGCTCGCCATTGGCGCAGTTCCGTCCGGCCTGCTCGCGCTCCTGGTGTTCTGCAT 654

Qy        695 GCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCT 754
```

```
                 |||  |||||||  ||||||||||||  |||||||  ||||||||||||||||  ||   |||||
Db        655 GCCCGAGTCGCCTCGGTGGCTGGTCTTGAAGGGCCGCCTCGCGGACGCCAGGGCTGTGCT 714

Qy        755 GGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGC 814
              ||||||||||  |  ||||  |  ||||||||||  |||||  ||||||||||||||||||||
Db        715 AGAGAAGACCTCTGCCACGCCAGAGGAGGCCGCCGAGCGGCTGGCCGACATCAAGGCCGC 774

Qy        815 CGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCC---CAAGAGAGGGAG 871
              ||  ||  ||  ||  |||  ||||||||  ||||| ||  |||||  ||  ||||         |
Db        775 GGCGGGGATTCCGAAGGGCCTCGACGGGGACGTAGTCACCGTACCCGGCAAGGAGCAAGG 834

Qy        872 CGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCG 931
              |||   |||  ||  ||||||||||  ||||||||||||||||||||||||  |  ||  ||
Db        835 CGGCGGTGAGTTGCAGGTGTGGAAGAAGCTCATCCTGTCCCCGACCCCGGCTGTCCGACG 894

Qy        932 CATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGT 991
              |||  |||||  ||  |   |  ||  ||||||||||||||||  ||  |  ||||   |||||||
Db        895 CATACTGCTCTCGGCCGTGGGTCTCCACTTCTTCCAGCAGGCTTCTGGCAGCGACTCCGT 954

Qy        992 CGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTT 1051
              ||||   ||||||  |  |  |  ||||||||||  |  ||  |  ||  |||||||  |  ||  |
Db        955 CGTCCAGTACAGCGCCCGCCTGTTCAAGAGCGCGGGGATCACCGACGACAACAAGCTCCT 1014

Qy       1052 GGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTT 1111
              ||||  |||  ||  ||  |  ||  ||  |||||||  |  |  ||||||  ||  ||||  ||  |||  |
Db       1015 GGGCGTCACCTGCGCGGTGGGCGTGACCAAGACGTTCTTCATCCTGGTGGCCACGTTCCT 1074

Qy       1112 CATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTC 1171
                 |  |||  |||  |||||  |||||  |||  |||||  |||||||||||||||  ||  ||||
Db       1075 GCTGGACCGCGCGGGGCGTCGGCCTCTGCTGCTGATCAGCACGGGCGGGATGATTGTCTC 1134

Qy       1172 CCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACC 1231
              ||||||  |||||||  |  ||||||||||||||  ||  ||  |||||  |||  |||||  |  |
Db       1135 GCTCATCTGCCTCGGGTCGGGGCTCACCGTCGCGGGGCATCACCCGGACACCAAGGTCGC 1194

Qy       1232 TTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGG 1291
              ||||||  |||  |||  |||||||  ||  |||||  |||||  ||||||||||||||||||
Db       1195 GTGGGCCGTCGCCCTGTGCATCGCGTCAACCCTGTCCTACATCGCCTTCTTCTCCATCGG 1254

Qy       1292 CCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCT 1351
              |||  ||  |||||||||  |  |||||||  ||||||  ||  ||||||||  ||||||||||||||
Db       1255 CCTCGGGCCCATCACGGGCGTGTACACCTCGGAAATATTCCCGCTGCAGGTGCGCGCGCT 1314

Qy       1352 GGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTT 1411
              |||||  |  ||  |  ||  ||  ||   ||||||||||||||||  ||||||||||||||||
Db       1315 GGGCTTCGCGGTGGGTGTGGCGAGCAACCGCGTCACCAGCGCCGTCATCTCCATGACCTT 1374

Qy       1412 CCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGC 1471
              ||||||  ||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1375 CCTGTCCCTCTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGC 1434

Qy       1472 CGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGA 1531
              ||||  ||||  |||||  |||||||  |||  |  ||||||||||||  ||||||||||||||||
Db       1435 CGCGGTCGCTTGGGTTTTCTTCTTCACGTGCCTCCCGGAGACACGCGGCCGGACGCTGGA 1494

Qy       1532 GGAGATGAGCAAGCTGTTCGGC------GACACGGCCGCCGCCTCGGAATCAGACGAGCC 1585
              |||||||  ||||||||||||||||      |||||||  |    ||     |||  ||||  ||  |
Db       1495 GGAGATGGGCAAGCTGTTCGGCATGCCAGACACGGGCATGGCTGAAGAAGCAGAAGACGC 1554

Qy       1586 AGCCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCACTAACT 1626
                 ||      ||  |||||  |||||||  ||  |    ||  ||  |
Db       1555 CGCAGCCAAGGAGAAGGTGGTGGAACTGCCTAGCAGCAAGT 1595
```

RESULT 6
ABX93207
ID ABX93207 standard; cDNA; 1914 BP.
XX
AC ABX93207;
XX
DT 29-MAY-2003 (first entry)
XX
DE cDNA encoding corn sugar transport protein #3.

XX
KW Beta vulgaris-like sugar transport protein; corn; rice; wheat;
KW plant sugar transport protein; carbohydrate transport; soybean;
KW carbohydrate distribution; plant; gene; ss.
XX
OS Zea mays.
XX
PN US2002178468-A1.
XX
PD 28-NOV-2002.
XX
PF 17-JAN-2002; 2002US-00051902.
XX
PR 24-APR-1998; 98US-0083044P.
PR 14-APR-1999; 99US-00291922.
XX
PA (ALLE/) ALLEN S M.
PA (HITZ/) HITZ W D.
PA (KINN/) KINNEY A J.
PA (TING/) TINGEY S V.
XX
PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX
DR WPI; 2003-340957/32.
DR P-PSDB; ABU08335.
XX
PT Novel plant sugar transport proteins and nucleic acid encoding the
PT protein useful for producing transgenic plants having altered levels of
PT sugar transport protein.
XX
PS Claim 7; Page 30; 56pp; English.
XX
CC The present invention relates to the isolation of Arabidopsis thaliana-
CC like or Beta vulgaris-like sugar transport proteins, and the
CC polynucleotide sequences encoding them. The plant sugar transport
CC proteins of the invention have been isolated from corn, rice, soybean,
CC and wheat. The polypeptides of the invention may be used for altering the
CC level of expression of a sugar transport protein in a host cell, by
CC transforming a host cell with a chimeric construct encoding all, or a
CC portion of the sugar transport protein, in sense or antisense
CC orientation. Particularly, the polypeptides may provide a means to
CC control carbohydrate transport and distribution in plants. ABX93206-
CC ABX93211 represent cDNA sequences encoding Beta vulgaris-like sugar
CC transport proteins
XX
SQ Sequence 1914 BP; 343 A; 591 C; 572 G; 408 T; 0 U; 0 Other;

Query Match 49.8%; Score 1005; DB 7; Length 1914;
Best Local Similarity 79.2%; Pred. No. 3.5e-162;
Matches 1220; Conservative 1; Mismatches 311; Indels 9; Gaps 2;

```
Qy         95 GATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTT 154
              |||||||||||  ||  |||   |  |||||||||  |||  | ||||||||||||||||   ||
Db         55 GATGGCTTCCGACGAGCTCGCAAAGGCCGTCGAGCCCAGGAAGAAGGGCAACGTCAAGTA 114

Qy        155 CGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGG 214
              ||||  |    ||  ||||||||  |||||||||  ||||   ||  ||||  |||||  ||  ||  ||
Db        115 TGCCTCCATATGTGCCATCCTGGCCTCCATGGCCTCTGTCATCCTTGGCTATGACATTGG 174

Qy        215 GGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGT 274
              |||||||||  ||  |||  |  ||||||||||||||||  |  ||  ||||   ||||  |  ||  |
Db        175 GGTGATGAGTGGAGCGGCCATGTACATCAAGAAGGACCTGAATATCACGGACGTGCAGCT 234

Qy        275 GGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCG 334
              ||||  |  ||  ||  ||  ||  ||  |  ||||||||||   ||||  ||||||||  |   |  |||
Db        235 GGAGATCCTGATCGGGATCCTCAGTCTCTACTCGCTGTTCGGATCCTTCGCTGGCGCGCG 294

Qy        335 GACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGC 394
              ||||||  |||  ||||||||  ||  ||  |  |||  ||||||||||||||  |||||  |||||||
Db        295 GACGTCCGACAGGATCGGGCGCCGCTTGACCGTCGTGTTCGCCGCTGTCATCTTCTTCGT 354

Qy        395 GGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGT 454
                 |||  :|||  ||||||||  ||||||||||||||||  ||||||||||   ||||||||||
```

```
Db        355 GGGCTCGTTGCTCATGGGTTTCGCCGTCAACTACGGCATGCTCATGGCGGGCCGCTTCGT 414

Qy        455 GGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTC 514
              ||||||  ||||  ||||||||||  |  ||||||||||||  ||||||||  ||||||  |  ||
Db        415 GGCCGGAGTCGGTGTGGGCTACGGGGGCATGATCGCGCCCGTGTACACGGCCGAGATCTC 474

Qy        515 GCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCAT 574
              |||  |||  ||  |  ||||||||||||||  |  ||||||||||||||||||||  |||||||
Db        475 GCCTGCGGCGTCCCGTGGCTTCCTGACCACCTTCCCGGAGGTGTTCATCAACATCGGCAT 534

Qy        575 CCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCG 634
              ||||||  ||  |||  |  ||  ||||  ||  |||  |  |||  |  |||||  |||||||  |||||
Db        535 CCTGCTTGGCTACCTGTCCAACTTCGCGTTCGCGCGCCTCCCGCTCCACCTCGGCTGGCG 594

Qy        635 CATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCAT 694
              |  ||||||||||  |||  |||||  |  |||||||  ||||||||||||  |||||  ||  ||||
Db        595 CGTCATGCTCGCCATTGGCGCAGTTCCGTCCGGCCTGCTCGCGCTCCTGGTGTTCTGCAT 654

Qy        695 GCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCT 754
              |||  ||||||||  ||||||||||||  |||||||  ||||||||||||||||  ||  |||||
Db        655 GCCCGAGTCGCCTCGGTGGCTGGTCTTGAAGGGCCGCCTCGCGGACGCCAGGGCTGTGCT 714

Qy        755 GGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGC 814
              ||||||||||  |  ||||  |  ||||||||||  |||||  ||||||||||||||||||
Db        715 AGAGAAGACCTCTGCCACGCCAGAGGAGGCCGCCGAGCGGCTGGCCGACATCAAGGCCGC 774

Qy        815 CGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCC---CAAGAGAGGGAG 871
              ||  ||  ||  ||  |||  ||||||||  |||||  ||  |||||  ||  ||||  |
Db        775 GGCGGGGATTCCGAAGGGCCTCGACGGGGACGTAGTCACCGTACCCGGCAAGGAGCAAGG 834

Qy        872 CGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCG 931
              |||  |||  ||  ||||||||||  ||||||||||||||||||||||||  |  ||  ||
Db        835 CGGCGGTGAGTTGCAGGTGTGGAAGAAGCTCATCCTGTCCCCGACCCCGGCTGTCCGACG 894

Qy        932 CATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGT 991
              |||  |||||  ||  |  |  ||  ||||||||||||||||  ||  |  ||||  |||||||
Db        895 CATACTGCTCTCGGCCGTGGGTCTCCACTTCTTCCAGCAGGCTTCTGGCAGCGACTCCGT 954

Qy        992 CGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTT 1051
              ||||  ||||||  |  |  |  ||||||||||  |  ||  |  ||  ||||||  |  ||  |
Db        955 CGTCCAGTACAGCGCCCGCCTGTTCAAGAGCGCGGGGATCACCGACGACAACAAGCTCCT 1014

Qy       1052 GGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTT 1111
              ||||  |||  ||  ||  |  ||  ||  |||||||  |  |  ||||||  ||  ||||  ||  |||  |
Db       1015 GGGCGTCACCTGCGCGGTGGGCGTGACCAAGACGTTCTTCATCCTGGTGGCCACGTTCCT 1074

Qy       1112 CATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGGATAATCCTCTC 1171
                 |  |||  |||  |||||  |||||  |||  |||||  ||||||||||||||  ||  ||||
Db       1075 GCTGGACCGCGCGGGGGCGTCGGCCTCTGCTGCTGATCAGCACGGGCGGGGATGATTGTCTC 1134

Qy       1172 CCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACC 1231
              ||||||  ||||||||  |  ||||||||||||  ||  ||  |||||  |||  |||||  |  |
Db       1135 GCTCATCTGCCTCGGGTCGGGGCTCACCGTCGCGGGGCATCACCCGGACACCAAGGTCGC 1194

Qy       1232 TTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGG 1291
              ||||||  |||  |||  |||||||  ||  |||||  |||||  ||||||||||||||||||
Db       1195 GTGGGCCGTCGCCCTGTGCATCGCGTCAACCCTGTCCTACATCGCCTTCTTCTCCATCGG 1254

Qy       1292 CCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCT 1351
              |||  ||  ||||||||  |  |||||||  ||||||  ||  ||||||||  ||||||||||||||
Db       1255 CCTCGGGCCCATCACGGGCGTGTACACCTCGGAAATATTCCCGCTGCAGGTGCGCGCGCT 1314

Qy       1352 GGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTT 1411
              ||||||  |  ||  |  ||  ||  ||  ||||||||||||||||  ||||||||||||||||||
Db       1315 GGGCTTCGCGGTGGGTGTGGCGAGCAACCGCGTCACCAGCGCCGTCATCTCCATGACCTT 1374

Qy       1412 CCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGC 1471
              ||||||  ||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1375 CCTGTCCCTCTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGC 1434

Qy       1472 CGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGA 1531
              ||||  ||||  |||||  ||||||  |||  |  ||||||||||  ||||||||||||||||||
Db       1435 CGCGGTCGCTTGGGTTTTCTTCTTCACGTGCCTCCCGGAGACACGCGGCCGGACGCTGGA 1494
```

```
Qy       1532 GGAGATGAGCAAGCTGTTCGGC------GACACGGCCGCCGCCTCGGAATCAGACGAGCC 1585
             ||||||| |||||||||||||||      ||||||| |   ||    ||| |||| ||  |
Db       1495 GGAGATGGGCAAGCTGTTCGGCATGCCAGACACGGGCATGGCTGAAGAAGCAGAAGACGC 1554

Qy       1586 AGCCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCACTAACT 1626
               ||      || |||||  ||||||| || |   ||  || |
Db       1555 CGCAGCCAAGGAGAAGGTGGTGGAACTGCCTAGCAGCAAGT 1595


RESULT 7
ABK51975
ID   ABK51975 standard; cDNA; 1872 BP.
XX
AC   ABK51975;
XX
DT   27-AUG-2002  (first entry)
XX
DE   Wheat cDNA clone wlm1.pk0012.h1 encoding sugar transport protein.
XX
KW   Wheat; Beta vulgaris-like sugar transport protein;
KW   carbohydrate transport; grain filling; annual field crop; plant;
KW   clone wlm1.pk0012.h1; gene; ss.
XX
OS   Triticum aestivum.
XX
FH   Key             Location/Qualifiers
FT   CDS             65. .1654
FT                   /*tag=  a
FT                   /product= "Wheat Beta vulgaris-like sugar transport
FT                   protein"
XX
PN   US6383776-B1.
XX
PD   07-MAY-2002.
XX
PF   14-APR-1999;   99US-00291922.
XX
PR   24-APR-1998;   98US-0083044P.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Allen SM,  Hitz WD,  Kinney AJ,  Tingey SV;
XX
DR   WPI; 2002-453364/48.
DR   P-PSDB; AAU97214.
XX
PT   New nucleic acid encoding plant sugar-transport proteins, useful for
PT   preparing transgenic plants with altered carbohydrate distribution.
XX
PS   Example 4; Col 75-78; 54pp; English.
XX
CC   The present invention relates to the isolation of plant polynucleotide
CC   sequences encoding an Arabidopsis thaliana-like sugar transport protein
CC   or Beta vulgaris-like sugar transport protein. The polynucleotide
CC   sequences are useful for altering the level of sugar transport proteins
CC   in plants, i.e. for control of carbohydrate transport and distribution in
CC   plant cells, e.g. during grain filling of annual field crops (e.g. corn,
CC   rice, soybeans, and wheat), and, for studying carbohydrate flows and
CC   sugar transport. The polynucleotide sequences can also be used to isolate
CC   cDNA sequences and genes that encode homologues of the new proteins. The
CC   present sequence encodes a wheat Beta vulgaris-like sugar transport
CC   protein
XX
SQ   Sequence 1872 BP; 390 A; 579 C; 534 G; 369 T; 0 U; 0 Other;

  Query Match             48.7%;  Score 981.6;  DB 6;  Length 1872;
  Best Local Similarity   77.4%;  Pred. No. 3.4e-158;
  Matches 1188;  Conservative    1;  Mismatches  345;  Indels    0;  Gaps    0;

Qy         99 GCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCC 158
              |||  |    | ||  ||||| ||| |||   || | ||| |||||||| ||  || ||||
Db        128 GCTGCGCTCCCGGAGCCGGGGGCAGTCCATCCAAGGAACAAGGGCAATTTCAAGTACGCC 187

Qy        159 TTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTG 218
              ||| |||||||| ||   || |||||| || ||||| |||||||||||||  | ||||||
Db        188 TTCACCTGCGCCCTCTGTGCTTCCATGGCCACCATCGTCCTCGGCTACGACGTTGGGGTG 247

Qy        219 ATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAG 278
              |||||||| |||||||||||||||||| |||| |  | ||||  |||| | || |||||
Db        248 ATGAGCGGTGCGTCGCTGTACATCAAGAGGGACCTGCAGATCACGGACGTGCAGCTGGAG 307

Qy        279 GTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACG 338
               |  | |||||||| |||| | | ||| |||||||||| ||||||  |  | | |||||
Db        308 ATCATGATGGGCATCCTGAGCGTGTACGCGCTCATCGGGTCCTTCCTCGGCGCGAGGACG 367

Qy        339 TCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGG 398
              || |||||| ||||||| ||   |||| |||| ||||| |||| ||| ||| |  ||
Db        368 TCCGACTGGGTCGGCCGCCGCGTCACCGTCGTCTTCGCGGCCGCCATCTTCAACAACGGC 427

Qy        399 GSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCC 458
               : || |||||||| ||||| |||||||||||||||||||| |||| |||||||  ||
Db        428 TCCTTGCTCATGGGCTTCGCGGTCAACTACGCCATGCTCATGGTCGGGCGCTTCGTCACC 487

Qy        459 GGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCG 518
              || |||||||||||||||  ||||| |||||| ||||||||  |||||||||  |||
Db        488 GGAATCGGCGTGGGCTACGCCATCATGGTCGCGCCAGTGTACACGCCCGAGGTGTCCCCG 547

Qy        519 GCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTG 578
              |||||||| || |||||||| |||||  ||| | ||||||||||||||  | |||||||||
Db        548 GCGTCGGCCCGCGGCTTCCTCACGTCTTTCACCGAGGTGTTCATCAATGTGGGCATCCTC 607

Qy        579 CTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATC 638
              || || |||||||| ||||| || ||| | ||| | |||||  |||||  | ||||||  ||
Db        608 CTTGGCTACGTCTCCAACTACGCCTTCGCGCGCCTCCCGCTCCACCTCAGCTGGCGCGTC 667

Qy        639 ATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCG 698
              |||||||||||||||| |  |||||||  |||||  |||||||||||| ||||||||||
Db        668 ATGCTCGGCATCGGCGCCGTCCCGTCCGCCCTGCTTGCGCTCATGGTGTTCGGCATGCCG 727

Qy        699 GAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAG 758
              ||||| || || |||||  |||||||| || ||||||||||||||  ||  || ||||
Db        728 GAGTCTCCTCGCTGGCTCGTCATGAAAGGCCGCCTCGCGGACGCCAGGGCCGTTCTGGCC 787

Qy        759 AAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCC 818
              |||||||||||||| |||||||||| |||||||| | | | |||||||| ||||||
Db        788 AAGACCTCCGACACGCCGGAGGAGGCCGTGGAGCGCCTTGACCAGATCAAGGCTGCCGCC 847

Qy        819 GGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAAC 878
              |||||||||  ||| ||  |||||||||||||||    | | ||  |||| |    |||| |
Db        848 GGCATCCCTAGGGAACTTGACGGCGACGTGGTCGTCATGCCTAAGACAAAAGGCGGCCAG 907

Qy        879 GAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTG 938
              ||||||| |||||||||||||||||| | || |||||||| |||||||||||||  |||
Db        908 GAGAAGCAGGTGTGGAAGGAGCTCATCTTTTCGCCGACCCCAGCCATGCGGCGCATACTG 967

Qy        939 CTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTC 998
              ||  | | | |||||||||| ||||| |||||  |||  ||||    |||||||||  ||
Db        968 CTCGCGGCGCTCGGCATCCATTTCTTTCAGCAGGCGACGGGCTCCGACTCCGTCGTGCTC 1027

Qy        999 TACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGCACC 1058
              || ||||| | ||||||| ||||| | ||  | ||   |||||| ||| |  | ||| ||
Db       1028 TATAGCCCACGCGTGTTCCAGAGCGCGGGCATCACCGGCGACAACCACCTGCTCGGCGCC 1087

Qy       1059 ACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGAC 1118
              || ||  |  | || ||||  |||| ||| |||||| || |||| || |||     |||||
Db       1088 ACATGCGCCATGGGGGTCATGAAGACGCTCTTCATCCTGGTGGCCACGTTCCAGCTCGAC 1147

Qy       1119 GGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATC 1178
              |||||||  |||||||||| |||||  |||||||  |||  ||  ||  ||  ||||||
Db       1148 CGCGTCGGCAGGCGGCCGCTGCTGCTGACCAGCACGGCCGGCATGCTCGCCTGTCTCATC 1207

Qy       1179 GGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCC 1238
              ||||||||  | || |||||||||| || | ||||||  |||||||| |  || ||||||
Db       1208 GGCCTCGGGACGGGCCTCACCGTCGTGGGTCGGCACCCGGACGCCAAGGTCCCGTGGGCC 1267
```

```
Qy       1239 ATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGC 1298
             ||||||||  ||||||  |||| | | ||||||||  ||||||||||||||||||||  ||
Db       1268 ATCGGCCTGTGCATCGTGTCCATCTTGGCCTACGTGTCCTTCTTCTCCATCGGCCTCGGG 1327

Qy       1299 CCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGC 1358
             |||  ||||  | |||||||  |||||||  |||||||  ||  | ||||||||||||||||  |
Db       1328 CCCCTCACCAGCGTGTACACCTCGGAGGTCTTCCCACTGCGGGTGCGCGCGCTGGGCTTC 1387

Qy       1359 TCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCG 1418
              ||||  |||   |   |||||||||||||||||  ||   ||||||||  ||||||||||
Db       1388 GCGCTGGGCACGTCATGCAACCGCGTCACCAGCGCCGCGGTCTCCATGTCCTTCCTGTCC 1447

Qy       1419 CTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTC 1478
             |||||||||||||||||||||||||||||||||||||||||||||||| ||| ||||||||||  |||  |
Db       1448 TTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTGTACGCCGGCATCGCGGCGATA 1507

Qy       1479 GCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATG 1538
             |  |||  | |||||||  |||||  |  | |||||||||  ||  ||||  |  ||||  |||||||||
Db       1508 GGATGGATTTTCTTCTTCACCTTCATTCCGGAGACGCGTGGCCTGCCGCTCGAGGAGATA 1567

Qy       1539 AGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAGCCAAGGAGAAG 1598
              |·|||||  ||||||   ||||  |  |  |||   |||  |   |  ||   |  ||||  |   |
Db       1568 GGGAAGCTTTTCGGCATGACGGACACGGCCGTCGAAGCCCAAGACACCGCCACGAAAGAC 1627

Qy       1599 AAGAAGGTGGAAATGGCCGCCACTAACTGATCAA 1632
              |||  |   |  |  ||     |   | |||  |  |
Db       1628 AAGGCGAAAGTAGGGGAGATGAACTAGTGAGCTA 1661
```

RESULT 8
ABX93211
ID ABX93211 standard; cDNA; 1872 BP.
XX
AC ABX93211;
XX
DT 29-MAY-2003 (first entry)
XX
DE cDNA encoding wheat sugar transport protein #5.
XX
KW Beta vulgaris-like sugar transport protein; corn; rice; wheat;
KW plant sugar transport protein; carbohydrate transport; soybean;
KW carbohydrate distribution; plant; gene; ss.
XX
OS Triticum aestivum.
XX
PN US2002178468-A1.
XX
PD 28-NOV-2002.
XX
PF 17-JAN-2002; 2002US-00051902.
XX
PR 24-APR-1998; 98US-0083044P.
PR 14-APR-1999; 99US-00291922.
XX
PA (ALLE/) ALLEN S M.
PA (HITZ/) HITZ W D.
PA (KINN/) KINNEY A J.
PA (TING/) TINGEY S V.
XX
PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX
DR WPI; 2003-340957/32.
DR P-PSDB; ABU08339.
XX
PT Novel plant sugar transport proteins and nucleic acid encoding the
PT protein useful for producing transgenic plants having altered levels of
PT sugar transport protein.
XX
PS Claim 7; Page 39-40; 56pp; English.
XX
CC The present invention relates to the isolation of Arabidopsis thaliana-
CC like or Beta vulgaris-like sugar transport proteins, and the
CC polynucleotide sequences encoding them. The plant sugar transport
CC proteins of the invention have been isolated from corn, rice, soybean,
CC and wheat. The polypeptides of the invention may be used for altering the
CC level of expression of a sugar transport protein in a host cell, by
CC transforming a host cell with a chimeric construct encoding all, or a
CC portion of the sugar transport protein, in sense or antisense
CC orientation. Particularly, the polypeptides may provide a means to
CC control carbohydrate transport and distribution in plants. ABX93206-
CC ABX93211 represent cDNA sequences encoding Beta vulgaris-like sugar
CC transport proteins
XX
SQ Sequence 1872 BP; 390 A; 579 C; 534 G; 369 T; 0 U; 0 Other;

Query Match 48.7%; Score 981.6; DB 7; Length 1872;
Best Local Similarity 77.4%; Pred. No. 3.4e-158;
Matches 1188; Conservative 1; Mismatches 345; Indels 0; Gaps 0;

```
Qy         99 GCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCC 158
              |||  |   | ||  |||||  |||  |||   ||  |  |||  ||||||||   ||  || ||||
Db        128 GCTGCGCTCCCGGAGCCGGGGGCAGTCCATCCAAGGAACAAGGGCAATTTCAAGTACGCC 187

Qy        159 TTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTG 218
              |||  |||||||| ||    ||  ||||||  ||  |||||  ||||||||||||||  |  |||||||
Db        188 TTCACCTGCGCCCTCTGTGCTTCCATGGCCACCATCGTCCTCGGCTACGACGTTGGGGTG 247

Qy        219 ATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAG 278
              ||||||||  ||||||||||||||||||||  ||||  |  |||||  ||||  |  ||  ||||||
Db        248 ATGAGCGGTGCGTCGCTGTACATCAAGAGGGACCTGCAGATCACGGACGTGCAGCTGGAG 307

Qy        279 GTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACG 338
                |  | ||||||||  ||||  |  |||  ||||||||||  ||||||    |  |  |  ||||||
Db        308 ATCATGATGGGCATCCTGAGCGTGTACGCGCTCATCGGGTCCTTCCTCGGCGCGAGGACG 367

Qy        339 TCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGG 398
              ||  ||||||  |||||||  ||    ||||  ||||  |||||  ||||  |||  |||  |    ||
Db        368 TCCGACTGGGTCGGCCGCCGCGTCACCGTCGTCTTCGCGGCCGCCATCTTCAACAACGGC 427

Qy        399 GSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCC 458
                :  ||  ||||||||  |||||  ||||||||||||||||||||  ||||  ||||||||  ||
Db        428 TCCTTGCTCATGGGCTTCGCGGTCAACTACGCCATGCTCATGGTCGGGCGCTTCGTCACC 487

Qy        459 GGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCG 518
              ||  ||||||||||||||||  |||||  |||||||  ||||||||  ||||||||||  |||
Db        488 GGAATCGGCGTGGGCTACGCCATCATGGTCGCGCCAGTGTACACGCCCGAGGTGTCCCCG 547

Qy        519 GCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTG 578
              ||||||||  ||  ||||||||  |||||  |||  | ||||||||||||||  |  |||||||||
Db        548 GCGTCGGCCCGCGGCTTCCTCACGTCTTTCACCGAGGTGTTCATCAATGTGGGCATCCTC 607

Qy        579 CTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATC 638
              ||  ||  ||||||||  |||||  ||  |||  |  |||  |  |||||  |||||  |  ||||||  ||
Db        608 CTTGGCTACGTCTCCAACTACGCCTTCGCGCGCCTCCCGCTCCACCTCAGCTGGCGCGTC 667

Qy        639 ATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCG 698
              ||||||||||||||||  |  |||||||  |||||  ||||||||||||  ||||||||||||
Db        668 ATGCTCGGCATCGGCGCCGTCCCGTCCGCCCTGCTTGCGCTCATGGTGTTCGGCATGCCG 727

Qy        699 GAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAG 758
              |||||  ||  ||  |||||  ||||||||  ||  ||||||||||||||||  ||   ||  ||||
Db        728 GAGTCTCCTCGCTGGCTCGTCATGAAAGGCCGCCTCGCGGACGCCAGGGCCGTTCTGGCC 787

Qy        759 AAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCC 818
              ||||||||||||||  ||||||||||  |||||||||  |  |  |  ||||||||  ||||||
Db        788 AAGACCTCCGACACGCCGGAGGAGGCCGTGGAGCGCCTTGACCAGATCAAGGCTGCCGCC 847

Qy        819 GGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAAC 878
              |||||||||  |||  ||  ||||||||||||||||  |  |  |||  ||||  |  |
Db        848 GGCATCCCTAGGGAACTTGACGGCGACGTGGTCGTCATGCCTAAGACAAAAGGCGGCCAG 907

Qy        879 GAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTG 938
              |||||||  ||||||||||||||||||  |  ||  ||||||||  ||||||||||||||  |||
Db        908 GAGAAGCAGGTGTGGAAGGAGCTCATCTTTTCGCCGACCCCAGCCATGCGGCGCATACTG 967

Qy        939 CTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTC 998
```

```
Db        968 CTCGCGGCGCTCGGCATCCATTTCTTTCAGCAGGCGACGGGCTCCGACTCCGTCGTGCTC 1027

Qy        999 TACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGCACC 1058

Db       1028 TATAGCCCACGCGTGTTCCAGAGCGCGGGCATCACCGGCGACAACCACCTGCTCGGCGCC 1087

Qy       1059 ACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGAC 1118

Db       1088 ACATGCGCCATGGGGGTCATGAAGACGCTCTTCATCCTGGTGGCCACGTTCCAGCTCGAC 1147

Qy       1119 GGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATC 1178

Db       1148 CGCGTCGGCAGGCGGCCGCTGCTGCTGACCAGCACGGCCGGCATGCTCGCCTGTCTCATC 1207

Qy       1179 GGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCC 1238

Db       1208 GGCCTCGGGACGGGCCTCACCGTCGTGGGTCGGCACCCGGACGCCAAGGTCCCGTGGGCC 1267

Qy       1239 ATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGC 1298

Db       1268 ATCGGCCTGTGCATCGTGTCCATCTTGGCCTACGTGTCCTTCTTCTCCATCGGCCTCGGG 1327

Qy       1299 CCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGC 1358

Db       1328 CCCCTCACCAGCGTGTACACCTCGGAGGTCTTCCCACTGCGGGTGCGCGCGCTGGGCTTC 1387

Qy       1359 TCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCG 1418

Db       1388 GCGCTGGGCACGTCATGCAACCGCGTCACCAGCGCCGCGGTCTCCATGTCCTTCCTGTCC 1447

Qy       1419 CTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTC 1478

Db       1448 TTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTGTACGCCGGCATCGCGGCGATA 1507

Qy       1479 GCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATG 1538

Db       1508 GGATGGATTTTCTTCTTCACCTTCATTCCGGAGACGCGTGGCCTGCCGCTCGAGGAGATA 1567

Qy       1539 AGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAGCCAAGGAGAAG 1598

Db       1568 GGGAAGCTTTTCGGCATGACGGACACGGCCGTCGAAGCCCAAGACACCGCCACGAAAGAC 1627

Qy       1599 AAGAAGGTGGAAATGGCCGCCACTAACTGATCAA 1632

Db       1628 AAGGCGAAAGTAGGGGAGATGAACTAGTGAGCTA 1661
```

```
RESULT 9
ABK51973
ID   ABK51973 standard; cDNA; 1853 BP.
XX
AC   ABK51973;
XX
DT   27-AUG-2002  (first entry)
XX
DE   Soybean contig encoding Beta vulgaris-like sugar transport protein.
XX
KW   Soybean; Beta vulgaris-like sugar transport protein;
KW   carbohydrate transport; grain filling; annual field crop; plant; gene;
KW   ss.
XX
OS   Glycine max.
XX
FH   Key             Location/Qualifiers
FT   CDS             167. .1738
FT                   /*tag=  a
FT                   /product= "Soybean Beta vulgaris-like sugar transport
FT                   protein"
XX
PN   US6383776-B1.
XX
PD   07-MAY-2002.
XX
PF   14-APR-1999;   99US-00291922.
XX
PR   24-APR-1998;   98US-0083044P.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Allen SM,  Hitz WD,  Kinney AJ,  Tingey SV;
XX
DR   WPI; 2002-453364/48.
DR   P-PSDB; AAU97212.
XX
PT   New nucleic acid encoding plant sugar-transport proteins, useful for
PT   preparing transgenic plants with altered carbohydrate distribution.
XX
PS   Example 4; Col 65-68; 54pp; English.
XX
CC   The present invention relates to the isolation of plant polynucleotide
CC   sequences encoding an Arabidopsis thaliana-like sugar transport protein
CC   or Beta vulgaris-like sugar transport protein. The polynucleotide
CC   sequences are useful for altering the level of sugar transport proteins
CC   in plants, i.e. for control of carbohydrate transport and distribution in
CC   plant cells, e.g. during grain filling of annual field crops (e.g. corn,
CC   rice, soybeans, and wheat), and, for studying carbohydrate flows and
CC   sugar transport. The polynucleotide sequences can also be used to isolate
CC   cDNA sequences and genes that encode homologues of the new proteins. The
CC   present sequence representing a contig assembled from various soybean
CC   cDNA clones encodes a Beta vulgaris-like sugar transport protein
XX
SQ   Sequence 1853 BP; 457 A; 434 C; 448 G; 513 T; 0 U; 1 Other;

  Query Match             28.5%;  Score 575.2;  DB 6;  Length 1853;
  Best Local Similarity   63.6%;  Pred. No. 6.7e-89;
  Matches  909;  Conservative    2;  Mismatches  510;  Indels    9;  Gaps    2;

Qy        109 CGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGCG 168

Db        210 CACTTCAGGATTTCGATCCTCCAAAGAAGCGCAAAAGGAACAAGTATGCTTTTGCTTGTG 269

Qy        169 CCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGGG 228

Db        270 CTATGCTGGCCTCCATGACTTCCATCTTGCTTGGTTATGATATTGGAGTGATGAGTGGAG 329

Qy        229 CGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGG 288

Db        330 CAGCCATATACATAAAAAGGGACCTGAAAGTCTCGGACGAGCAAATCGAGATCCTGCTCG 389

Qy        289 GCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGA 348

Db        390 GAATCATCAACCTATACTCTCTGATAGGCTCATGTCTCGCCGGCAGAACCTCCGACTGGA 449

Qy        349 TCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGSGTTCCTCA 408

Db        450 TAGGTCCCCGTTACACGATTGTTTTCGCCGGCACCATCTTCTTTGTCGGAGCACTTCTCA 509

Qy        409 TGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCG 468

Db        510 TGGGTTTCTCCCCCAATTATTCCTTTCTCATGTTTGGCCGTTTCGTCGCTGGCATTGGCA 569

Qy        469 TGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGC 528

Db        570 TCGGCTACGCCCTCATGATAGCCCCCGTCTACACCGCCGAGGTCTCCCCGGCCTCCTCTC 629

Qy        529 GTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACG 588

Db        630 GTGGCTTCCTCACTTCCTTCCCTGAGGTATTTATTAATGGAGGGATATTAATTGGATACA 689

Qy        589 TCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCA 648

Db        690 TATCAAACTATGCATTTTCGAAGCTGACACTAAAGGTGGGATGGCGAATGATGCTTGGAG 749

Qy        649 TCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGC 708
```

```
Db        750 TTGGTGCAATACCTTCGGTACTCCTAACAGTAGGAGTGTTGGCGATGCCGGAGTCCCCAA 809

Qy        709 GGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCG 768

Db        810 GGTGGCTTGTGATGAGGGGTCGTTTGGGAGAGGCAAGAAAAGTGCTTAACAAAACCTCAG 869

Qy        769 ACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTG 828

Db        870 ACAGCAAGGAAGAGGCCCAACTAAGGCTAGCGGAAATCAAACAAGCCGCAGGGATCCCCG 929

Qy        829 AGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGG 888

Db        930 AGAGTTGCAACGACGACGTCGTTCAGGTAAATAAACAAAGCAACGGTGAAG------GTG 983

Qy        889 TGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGA 948

Db        984 TATGGAAAGAGCTCTTCCTCTATCCAACGCCCGCAATTCGTCACATCGTAATCGCTGCCC 1043

Qy        949 TCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTC 1008

Db       1044 TTGGTATTCACTTCTTCCAACAAGCGTCGGGCGTAGACGCCGTCGTTTTGTACAGCCCCA 1103

Qy       1009 TCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCGT 1068

Db       1104 GGATCTTCGAAAAGGCTGGGATTACAAACGACACGCATAAGCTTCTTGCAACCGTGGCCG 1163

Qy       1069 TCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGC 1128

Db       1164 TTGGATTCGTTAAGACCGTGTTCATCTTGGCGGCTACGTTTACGTTGGACCGCGTGGGTC 1223

Qy       1129 GGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCG 1188

Db       1224 GTCGTCCGTTGTTATTGTCTAGTGTCGGCGGCATGGTGCTCTCGCTTCTCACGCTTGCGA 1283

Qy       1189 CCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAA 1248

Db       1284 TCAGCCTCACTGTTATTG---ATCATTCGGAGAGGAAATTAATGTGGGCCGTTGGATCGA 1340

Qy       1249 GCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGT 1308

Db       1341 GCATAGCCATGGTGTTGGCTTACGTGGCCACGTTCTCCATCGGTGCGGGTCCCATCACGT 1400

Qy       1309 GGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCG 1368

Db       1401 GGGTCTATAGTTCTGAGATCTTCCCGTTGAGGCTGCGGGCGCARGGTGCGGCCGCGGGAG 1460

Qy       1369 TCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGG 1428

Db       1461 TTGCGGTGAATAGGACCACTAGCGCGGTTGTCTCAATGACTTTTCTGTCCCTCACTAGAG 1520

Qy       1429 CCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGT 1488

Db       1521 CCATCACTATTGGTGGAGCTTTCTTCCTTTATTGTGGCATTGCTACTGTTGGGTGGATAT 1580

Qy       1489 TCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATG 1538

Db       1581 TCTTTTACACCGTCTTGCCTGAGACCCGGGGAAAAACGCTCGAAGACATG 1630


RESULT 10
ABX93209
ID   ABX93209 standard; cDNA; 1853 BP.
XX
AC   ABX93209;
XX
DT   29-MAY-2003  (first entry)
XX
DE   cDNA encoding soybean sugar transport protein #3.
XX
KW   Beta vulgaris-like sugar transport protein; corn; rice; wheat;
KW   plant sugar transport protein; carbohydrate transport; soybean;
KW   carbohydrate distribution; plant; gene; ss.
XX
OS   Glycine max.
XX
PN   US2002178468-A1.
XX
PD   28-NOV-2002.
XX
PF   17-JAN-2002; 2002US-00051902.
XX
PR   24-APR-1998;   98US-0083044P.
PR   14-APR-1999;   99US-00291922.
XX
PA   (ALLE/) ALLEN S M.
PA   (HITZ/) HITZ W D.
PA   (KINN/) KINNEY A J.
PA   (TING/) TINGEY S V.
XX
PI   Allen SM,  Hitz WD,  Kinney AJ,  Tingey SV;
XX
DR   WPI; 2003-340957/32.
DR   P-PSDB; ABU08337.
XX
PT   Novel plant sugar transport proteins and nucleic acid encoding the
PT   protein useful for producing transgenic plants having altered levels of
PT   sugar transport protein.
XX
PS   Claim 7; Page 34-35; 56pp; English.
XX
CC   The present invention relates to the isolation of Arabidopsis thaliana-
CC   like or Beta vulgaris-like sugar transport proteins, and the
CC   polynucleotide sequences encoding them. The plant sugar transport
CC   proteins of the invention have been isolated from corn, rice, soybean,
CC   and wheat. The polypeptides of the invention may be used for altering the
CC   level of expression of a sugar transport protein in a host cell, by
CC   transforming a host cell with a chimeric construct encoding all, or a
CC   portion of the sugar transport protein, in sense or antisense
CC   orientation. Particularly, the polypeptides may provide a means to
CC   control carbohydrate transport and distribution in plants. ABX93206-
CC   ABX93211 represent cDNA sequences encoding Beta vulgaris-like sugar
CC   transport proteins
XX
SQ   Sequence 1853 BP; 457 A; 434 C; 448 G; 513 T; 0 U; 1 Other;

  Query Match             28.5%;  Score 575.2;  DB 7;  Length 1853;
  Best Local Similarity   63.6%;  Pred. No. 6.7e-89;
  Matches  909;  Conservative    2;  Mismatches  510;  Indels    9;  Gaps    2;

Qy        109 CGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGCG 168

Db        210 CACTTCAGGATTTCGATCCTCCAAAGAAGCGCAAAAGGAACAAGTATGCTTTTGCTTGTG 269

Qy        169 CCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGGG 228

Db        270 CTATGCTGGCCTCCATGACTTCCATCTTGCTTGGTTATGATATTGGAGTGATGAGTGGAG 329

Qy        229 CGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGG 288

Db        330 CAGCCATATACATAAAAAGGGACCTGAAAGTCTCGGACGAGCAAATCGAGATCCTGCTCG 389

Qy        289 GCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGA 348

Db        390 GAATCATCAACCTATACTCTCTGATAGGCTCATGTCTCGCCGGCAGAACCTCCGACTGGA 449

Qy        349 TCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGGSGTTCCTCA 408

Db        450 TAGGTCCCCGTTACACGATTGTTTTCGCCGGCACCATCTTCTTTGTCGGAGCACTTCTCA 509

Qy        409 TGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCG 468

Db        510 TGGGTTTCTCCCCCAATTATTCCTTTCTCATGTTTGGCCGTTTCGTCGCTGGCATTGGCA 569

Qy        469 TGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGC 528

Db        570 TCGGCTACGCCCTCATGATAGCCCCCGTCTACACCGCCGAGGTCTCCCCGGCCTCCTCTC 629
```

```
Qy       529 GTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACG 588
Db       630 GTGGCTTCCTCACTTCCTTCCCTGAGGTATTTATTAATGGAGGGATATTAATTGGATACA 689

Qy       589 TCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCA 648
Db       690 TATCAAACTATGCATTTTCGAAGCTGACACTAAAGGTGGGATGGCGAATGATGCTTGGAG 749

Qy       649 TCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGC 708
Db       750 TTGGTGCAATACCTTCGGTACTCCTAACAGTAGGAGTGTTGGCGATGCCGGAGTCCCCAA 809

Qy       709 GGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCG 768
Db       810 GGTGGCTTGTGATGAGGGGTCGTTTGGGAGAGGCAAGAAAAGTGCTTAACAAAACCTCAG 869

Qy       769 ACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTG 828
Db       870 ACAGCAAGGAAGAGGCCCAACTAAGGCTAGCGGAAATCAAACAAGCCGCAGGGATCCCCG 929

Qy       829 AGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGG 888
Db       930 AGAGTTGCAACGACGACGTCGTTCAGGTAAATAAACAAAGCAACGGTGAAG------GTG 983

Qy       889 TGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGA 948
Db       984 TATGGAAAGAGCTCTTCCTCTATCCAACGCCCGCAATTCGTCACATCGTAATCGCTGCCC 1043

Qy       949 TCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTC 1008
Db      1044 TTGGTATTCACTTCTTCCAACAAGCGTCGGGCGTAGACGCCGTCGTTTTGTACAGCCCCA 1103

Qy       1009 TCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCGT 1068
Db       1104 GGATCTTCGAAAAGGCTGGGATTACAAACGACACGCATAAGCTTCTTGCAACCGTGGCCG 1163

Qy       1069 TCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGC 1128
Db       1164 TTGGATTCGTTAAGACCGTGTTCATCTTGGCGGCTACGTTTACGTTGGACCGCGTGGGTC 1223

Qy       1129 GGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCG 1188
Db       1224 GTCGTCCGTTGTTATTGTCTAGTGTCGGCGGCATGGTGCTCTCGCTTCTCACGCTTGCGA 1283

Qy       1189 CCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAA 1248
Db       1284 TCAGCCTCACTGTTATTG---ATCATTCGGAGAGGAAATTAATGTGGGCCGTTGGATCGA 1340

Qy       1249 GCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGT 1308
Db       1341 GCATAGCCATGGTGTTGGCTTACGTGGCCACGTTCTCCATCGGTGCGGGTCCCATCACGT 1400

Qy       1309 GGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCG 1368
Db       1401 GGGTCTATAGTTCTGAGATCTTCCCGTTGAGGCTGCGGGCGCARGGTGCGGCCGCGGGAG 1460

Qy       1369 TCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGG 1428
Db       1461 TTGCGGTGAATAGGACCACTAGCGCGGTTGTCTCAATGACTTTTCTGTCCCTCACTAGAG 1520

Qy       1429 CCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGT 1488
Db       1521 CCATCACTATTGGTGGAGCTTTCTTCCTTTATTGTGGCATTGCTACTGTTGGGTGGATAT 1580

Qy       1489 TCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATG 1538
Db       1581 TCTTTTACACCGTCTTGCCTGAGACCCGGGGAAAAACGCTCGAAGACATG 1630
```

RESULT 11
ABL41873
ID ABL41873 standard; DNA; 2056 BP.
XX
AC ABL41873;
XX
DT 11-JUN-2002 (first entry)
XX
DE Nucleotide sequence of a linear polyol transporter.
XX
KW Linear polyol; mannitol; polyol transporter; carbon source; plant;
KW pathogen resistance; salt stress; sorbitol; dulcitol; galactitol;
KW inositol; ribitol; xylitol; ss.
XX
OS Beta vulgaris.
XX
PN WO200204647-A1.
XX
PD 17-JAN-2002.
XX
PF 22-JUN-2001; 2001WO-FR001979.
XX
PR 11-JUL-2000; 2000FR-00009032.
XX
PA (CNRS) CENT NAT RECH SCI.
XX
PI Lemoine RRP, Noiraud NEJ;
XX
DR WPI; 2002-154933/20.
XX
PT New polyol transporter protein from plants, for selecting transformed
PT cells and for imparting pathogen and salt-stress resistance to plants.
XX
PS Disclosure; Page 41-42; 66pp; French.
XX
CC The present sequence encodes a polypeptide which is a transporter of a
CC linear polyol. DNA encoding linear polyols is used to produce a selection
CC system for transformed cells, based on the use of polyols as the only
CC carbon source. It is also used to produce transgenic plants with
CC increased resistance to pathogens and salt stress. The use of a linear
CC polyol for selection eliminates the need for toxic selection reagents
CC such as antibiotics. Also, the linear polyol is not essential for the
CC plant, once selection has been made. The linear polyol has a main chain
CC of 5-8, preferably 6, carbon atoms and is selected from sorbitol,
CC dulcitol, galactitol, (myo-)inositol, ribitol, or xylitol
XX
SQ Sequence 2056 BP; 570 A; 374 C; 444 G; 667 T; 0 U; 1 Other;

Query Match 23.7%; Score 477.2; DB 6; Length 2056;
Best Local Similarity 59.2%; Pred. No. 3.5e-72;
Matches 831; Conservative 1; Mismatches 569; Indels 3; Gaps 1;

```
Qy       152 GTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATAT 211
Db       198 GTTTGCTTTTGCTTGTGCTACTTTGGCTTCTATGACTTCTGTTTTACTTGGTTATGACAT 257

Qy       212 CGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAA 271
Db       258 TGGAGTAATGAGTGGTGCAATAATCTACCTCAAAGAAGACTGGCACATTAGTGACACACA 317

Qy       272 GGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGG 331
Db       318 AATAGGAGTTCTAGTCGGAATCTTAAACATCTACTGCCTCTTCGGCTCTTTCGCAGCTGG 377

Qy       332 GCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTT 391
Db       378 TCGAACTTCCGATTGGATCGGACGACGTTACACCATCGTATTAGCTGGTGCAATCTTCTT 437

Qy       392 CGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTT 451
Db       438 CGTTGGCGCACTACTCATGGGTTTCGCCACAAACTACGCATTTCTCATGGTAGGCCGTTT 497

Qy       452 CGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGT 511
Db       498 TGTAACAGGAATAGGTGTTGGTTATGCACTTATGATCGCACCTGTTTACACTGCAGAGGT 557
```

```
Qy        512 GTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGG 571
Db        558 TTCTCCTGCTTCTTCAAGAGGATTTCTTACCTCTTTTCCTGAGGTTTTCATCAATGCAGG 617

Qy        572 CATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTG 631
Db        618 AATTTTGCTTGGATATATATCTAACCTTGCATTTTCAAGCCTCCCAACTCATTTGAGCTG 677

Qy        632 GCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGG 691
Db        678 GAGGTTTATGCTCGGAATCGGGGCGATTCCGAGCATATTTTTGGCTATTGGAGTGCTTGC 737

Qy        692 CATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGT 751
Db        738 TATGCCTGAGTCACCGCGGTGGCTTGTTATGCAGGGAAGGCTTGGAGATGCTAAGAAGGT 797

Qy        752 GCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGC 811
Db        798 GCTTAATCGCATATCTGACTCGCCTGAGGAGGCTCAACTTAGGCTCAGTGAGATTAAGCA 857

Qy        812 CGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAG 871
Db        858 GACAGCAGGAATCCCAGCTGAGTGTGATGAAGACATCTATAAGGTTGAAAAAACAAAGAT 917

Qy        872 CGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCG 931
Db        918 AAAATCAGGGAATGCAGTTTGGAAAGAGCTCTTCTTCAACCCTACTCCCGCAGTGAGGCG 977

Qy        932 CATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGT 991
Db        978 TGCAGTGATTGCAGGTATCGGAATTCACTTTTTCCAGCAAGCTTCCGGCATCGACGCCGT 1037

Qy        992 CGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTT 1051
Db       1038 GGTTTTATACAGTCCAAGGATCTTTCAGAGTGCTGGAATCACAAATGCACGTAAACAGTT 1097

Qy       1052 GGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTT 1111
Db       1098 ACTTGCAACAGTAGCTGTAGGAGTTGTTAAGACACTTTTCATATTAGTTGCTACATTTCA 1157

Qy       1112 CATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTC 1171
Db       1158 ATTAGACAAGTATGGTAGAAGGCCTTTACTATTAACAAGTGTTGGTGGTATGATCATAGC 1217

Qy       1172 CCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACC 1231
Db       1218 CATCTTAACCTTAGCCATGTCTTTAACTGTTATTG---ATCACTCTCATCACAAGATTAC 1274

Qy       1232 TTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGG 1291
Db       1275 ATGGGCTATAGCCTTGTGTATAACCATGGTTTGTGCTGTCGTTGCGTCGTTTTCGATTGG 1334

Qy       1292 CCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCT 1351
Db       1335 TTTAGGACCAATTACATGGGTATATAGTTCAGAGGTTTTCCCCTTAAGGTTAAGGGCTCA 1394

Qy       1352 GGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTT 1411
Db       1395 AGGTACTAGCATGGGTGTGGCTGTTAATAGAGTTGTAAGTGGTGTTATTTCAATATTTTT 1454

Qy       1412 CCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGC 1471
Db       1455 CTTGCCTTTGTCACATAAGATTACAACAGGTGGTGCCTTCTTTTTGTTCGGAGGGATTGC 1514

Qy       1472 CGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGA 1531
Db       1515 TATTATTGCTTGGTTCTTCTTCTTGACGTTTCTTCCTGAGACTAGAGGGCGTACACTTGA 1574

Qy       1532 GGAGATGAGCAAGCTGTTCGGCGA 1555
Db       1575 GAATATGCATGAGTTGTTTGAAGA 1598
```

RESULT 12
AAC43261
ID AAC43261 standard; DNA; 1527 BP.
XX
AC AAC43261;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 38621.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.

PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 23.3%; Score 470.2; DB 3; Length 1527;
Best Local Similarity 59.4%; Pred. No. 5.2e-71;
Matches 834; Conservative 1; Mismatches 559; Indels 9; Gaps 2;

```
Qy      152 GTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATAT 211
            |||||| ||||  || |||||  | || ||||| |  || ||| || | || || ||||
Db       60 GTTCGCTTTCGGGTGTGCCATTGTTGCTTCCATCATTTCTATCATCTTTGGATATGATAC 119

Qy      212 CGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAA 271
             || || |||||||| ||   | | |  || |   | ||  | ||  || |  |||     |
Db      120 GGGAGTTATGAGCGGAGCTCAGATATTTATAAGAGATGATCTAAAAATAAACGACACTCA 179

Qy      272 GGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGG 331
            | | || ||| |    || ||  |||| || |   | ||| |||| ||  |  |||||||
Db      180 GATTGAAGTTTTGGCCGGAATCTTGAATCTTTGTGCACTCGTCGGATCACTAACGGCGGG 239
```

```
Qy       332 GCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTT 391
Db       240 GAAAACGTCTGACGTCATCGGCCGACGTTACACCATAGCTCTCTCCGCCGTGATATTCTT 299

Qy       392 CGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTT 451
Db       300 AGTGGGCTCAGTTCTTATGGGTTACGGCCCAAACTATCCTGTTTTGATGGTCGGACGATG 359

Qy       452 CGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGT 511
Db       360 TATTGCCGGAGTTGGTGTAGGGTTTGCTCTCATGATAGCTCCGGTTTACTCTGCCGAGAT 419

Qy       512 GTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGG 571
Db       420 ATCCTCTGCGTCACATAGAGGCTTTCTCACTTCTCTACCTGAGCTTTGTATTAGTCTTGG 479

Qy       572 CATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTG 631
Db       480 GATTTTACTAGGCTATGTCTCGAATTACTGTTTTGGGAAACTGACGTTGAAACTTGGATG 539

Qy       632 GCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGG 691
Db       540 GAGATTGATGCTTGGAATCGCAGCGTTCCCGTCTTTGATATTGGCTTTTGGGATTACAAG 599

Qy       692 CATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGT 751
Db       600 AATGCCCGAGTCGCCGAGGTGGCTTGTGATGCAAGGTAGGTTAGAAGAAGCCAAGAAAAT 659

Qy       752 GCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGC 811
Db       660 AATGGTTTTGGTATCCAATACGGAAGAAGAAGCTGAAGAACGGTTCAGAGACATCTTAAC 719

Qy       812 CGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAG 871
Db       720 CGCTGCGGAAGTAGATGTGACAGAGATAAAAGAAGTTGGTGGCGGCGTTAAGAAGAAGAA 779

Qy       872 CGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCG 931
Db       780 TC---ACGGTAAAAGTGTTTGGAGGGAGTTGGTGATAAAACCGCGACCTGCGGTGCGTTT 836

Qy       932 CATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGT 991
Db       837 GATCTTGATAGCGGCCGTTGGGATACACTTTTTCGAGCATGCGACGGGGATTGAGGCCGT 896

Qy       992 CGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTT 1051
Db       897 CGTGTTGTACAGCCCGAGAATTTTCAAGAAAGCCGGAGTTGTTTCGAAAGACAAGCTCTT 956

Qy      1052 GGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTT 1111
Db       957 ACTAGCCACGGTCGGTGTAGGTTTAACTAAGGCCTTTTTCATAATAATAGCCACTTTCTT 1016

Qy      1112 CATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTC 1171
Db      1017 GCTCGACAAGGTAGGTCGTAGAAAGCTCCTGTTGACCAGTACGGGTGGGATGGTCTTCGC 1076

Qy      1172 CCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACC 1231
Db      1077 ATTGACCAGTTTAGCGGTTAGTCTCACGATGGT------TCAGCGATTTGGGCGGTTAGC 1130

Qy      1232 TTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGG 1291
Db      1131 GTGGGCATTGAGTTTAAGCATCGTCTCTACGTATGCATTTGTAGCTTTTTTCTCCATAGG 1190

Qy      1292 CCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCT 1351
Db      1191 GCTTGGACCTATCACATGGGTGTACAGCTCCGAGATATTCCCGTTGAGACTTAGGGCTCA 1250

Qy      1352 GGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTT 1411
Db      1251 GGGAGCGAGTATAGGTGTTGCGGTTAATAGGATCATGAACGCGACGGTCTCAATGAGTTT 1310

Qy      1412 CCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGC 1471

Db      1311 TCTGTCGATGACGAAGGCGATCACGACAGGGGGCGTGTTCTTCGTATTCGCCGGAATAGC 1370

Qy      1472 CGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGA 1531
Db      1371 GGTGGCGGCATGGTGGTTCTTCTTCTTTATGTTGCCGGAGACGAAAGGATTGCCATTGGA 1430

Qy      1532 GGAGATGAGCAAGCTGTTCGGCG 1554
Db      1431 GGAGATGGAGAAGCTTTTTGGTG 1453
```

RESULT 13
ABL41874
ID ABL41874 standard; DNA; 1925 BP.
XX
AC ABL41874;
XX
DT 11-JUN-2002 (first entry)
XX
DE Nucleotide sequence of a linear polyol transporter.
XX
KW Linear polyol; mannitol; polyol transporter; carbon source; plant;
KW pathogen resistance; salt stress; sorbitol; dulcitol; galactitol;
KW inositol; ribitol; xylitol; ss.
XX
OS Beta vulgaris.
XX
PN WO200204647-A1.
XX
PD 17-JAN-2002.
XX
PF 22-JUN-2001; 2001WO-FR001979.
XX
PR 11-JUL-2000; 2000FR-00009032.
XX
PA (CNRS) CENT NAT RECH SCI.
XX
PI Lemoine RRP, Noiraud NEJ;
XX
DR WPI; 2002-154933/20.
XX
PT New polyol transporter protein from plants, for selecting transformed
PT cells and for imparting pathogen and salt-stress resistance to plants.
XX
PS Disclosure; Page 42; 66pp; French.
XX
CC The present sequence encodes a polypeptide which is a transporter of a
CC linear polyol. DNA encoding linear polyols is used to produce a selection
CC system for transformed cells, based on the use of polyols as the only
CC carbon source. It is also used to produce transgenic plants with
CC increased resistance to pathogens and salt stress. The use of a linear
CC polyol for selection eliminates the need for toxic selection reagents
CC such as antibiotics. Also, the linear polyol is not essential for the
CC plant, once selection has been made. The linear polyol has a main chain
CC of 5-8, preferably 6, carbon atoms and is selected from sorbitol,
CC dulcitol, galactitol, (myo-)inositol, ribitol, or xylitol
XX
SQ Sequence 1925 BP; 529 A; 349 C; 432 G; 615 T; 0 U; 0 Other;

Query Match 23.3%; Score 469.2; DB 6; Length 1925;
Best Local Similarity 58.8%; Pred. No. 8e-71;
Matches 826; Conservative 1; Mismatches 574; Indels 3; Gaps 1;

```
Qy       152 GTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATAT 211
Db       144 GTTTGCTTTTGCTTGTGCTACTTTGGCTTCTATGACTTCTGTTTTACTTGGTTATGACAT 203

Qy       212 CGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAA 271
Db       204 TGGAGTAATGAGTGGTGCAATAATCTACCTCAAAGAAGACTGGCACATTAGTGACACACA 263

Qy       272 GGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGG 331
```

```
Db        264 AATAGGAGTTCTAGTCGGAATCTTAAACATCTACTGCCTCTTCGGCTCTTTCGCAGCTGG 323

Qy        332 GCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTT 391

Db        324 TAGAACATCCGATTGGATCGGACGACGTTACACTATCGTGTTAGCTGGTGCAATCTTCTT 383

Qy        392 CGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTT 451

Db        384 CGTTGGCGCACTACTCATGGGTTTCGCCACAAACTACGCATTTCTTATGGTAGGCCGTTT 443

Qy        452 CGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGT 511

Db        444 TGTAACAGGAATAGGTGTTGGTTATGCACTTATGATCGCACCTGTTTACACTGCAGAGGT 503

Qy        512 GTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGG 571

Db        504 TTCTCCTGCTTCTTCAAGAGGATTTCTTACCTCTTTTCCTGAGGTTTTCATCAATGCAGG 563

Qy        572 CATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTG 631

Db        564 AATTTTGCTTGGATATATATCTAATCTCGCATTTTCAAGCCTCCCAACTCATTTGAGCTG 623

Qy        632 GCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGG 691

Db        624 GAGGTTTATGCTCGGAATCGGGGCGATTCCGAGCATATTTTTGGCTATTGGAGTGCTTGC 683

Qy        692 CATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGT 751

Db        684 TATGCCTGAGTCACCGAGGTGGCTTGTTATGCAGGGAAGGCTTGGAGATGCTAAGAAGGT 743

Qy        752 GCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGC 811

Db        744 GCTTAATCGAATATCTGACTCGCCTGAGGAGGCTCAACTTAGGCTCAGTGAGATTAAGCA 803

Qy        812 CGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAG 871

Db        804 GACAGCAGGAATCCCAGCTGAGTGTGATGAAGACATATATAAGGTTGAAAAAACAAAGAT 863

Qy        872 CGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCG 931

Db        864 AAAATCAGGGAATGCAGTTTGGAAAGAGCTCTTCTTCAACCCTACTCCCGCAGTGAGGCG 923

Qy        932 CATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGT 991

Db        924 TGCAGTGATTGCAGGTATCGGAATTCACTTTTTCCAGCAAGCTTCCGGCATCGACGCCGT 983

Qy        992 CGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTT 1051

Db        984 GGTTTTATACAGTCCAAGGATCTTTCAGAGTGCTGGAATCACAAATGCACGTAAACAGTT 1043

Qy       1052 GGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTT 1111

Db       1044 ACTTGCAACAGTAGCTGTAGGAGTTGTTAAGACACTTTTCATATTAGTTGCTACATTTCA 1103

Qy       1112 CATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTC 1171

Db       1104 ATTAGACAAGTATGGTAGAAGGCCTTTACTATTAACAAGTGTTGGTGGTATGATCATAGC 1163

Qy       1172 CCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACC 1231

Db       1164 CATCTTAACCTTAGCCATGTCTTTAACTGTTATTG---ATCACTCTCATCACAAGATTAC 1220

Qy       1232 TTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGG 1291

Db       1221 ATGGGCTATAGCCTTGTGTATAACCATGGTTTGTGCTGTCGTTGCGTCGTTTTCGATTGG 1280

Qy       1292 CCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCT 1351

Db       1281 TTTAGGACCAATTACATGGGTATATAGTTCAGAGGTTTTCCCCTTGAGGTTAAGGGCTCA 1340

Qy       1352 GGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTT 1411

Db       1341 AGGTACTAGCATGGGTGTGGCTGTTAATAGAGTTGTAAGTGGTGTTATTTCAATATTTTT 1400

Qy       1412 CCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGC 1471

Db       1401 CTTGCCTTTGTCACATAAGATTACAACAGGTGGTGCCTTCTTTTTGTTCGGAGGGATTGC 1460

Qy       1472 CGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGA 1531

Db       1461 TATTATTGCTTGGTTCTTCTTCTTGACGTTTCTTCCTGAGACTAGAGGGCGTACACTTGA 1520

Qy       1532 GGAGATGAGCAAGCTGTTCGGCGA 1555

Db       1521 GAATATGCATGAGTTGTTTGAAGA 1544
```

RESULT 14
ABL41878/c
ID ABL41878 standard; DNA; 1662 BP.
XX
AC ABL41878;
XX
DT 11-JUN-2002 (first entry)
XX
DE Nucleotide sequence of a linear polyol transporter.
XX
KW Linear polyol; mannitol; polyol transporter; carbon source; plant;
KW pathogen resistance; salt stress; sorbitol; dulcitol; galactitol;
KW inositol; ribitol; xylitol; ss.
XX
OS Arabidopsis thaliana.
XX
PN WO200204647-A1.
XX
PD 17-JAN-2002.
XX
PF 22-JUN-2001; 2001WO-FR001979.
XX
PR 11-JUL-2000; 2000FR-00009032.
XX
PA (CNRS) CENT NAT RECH SCI.
XX
PI Lemoine RRP, Noiraud NEJ;
XX
DR WPI; 2002-154933/20.
XX
PT New polyol transporter protein from plants, for selecting transformed
PT cells and for imparting pathogen and salt-stress resistance to plants.
XX
PS Disclosure; Page 44-45; 66pp; French.
XX
CC The present sequence encodes a polypeptide which is a transporter of a
CC linear polyol. DNA encoding linear polyols is used to produce a selection
CC system for transformed cells, based on the use of polyols as the only
CC carbon source. It is also used to produce transgenic plants with
CC increased resistance to pathogens and salt stress. The use of a linear
CC polyol for selection eliminates the need for toxic selection reagents
CC such as antibiotics. Also, the linear polyol is not essential for the
CC plant, once selection has been made. The linear polyol has a main chain
CC of 5-8, preferably 6, carbon atoms and is selected from sorbitol,
CC dulcitol, galactitol, (myo-)inositol, ribitol, or xylitol
XX
SQ Sequence 1662 BP; 508 A; 445 C; 306 G; 403 T; 0 U; 0 Other;

  Query Match   22.0%; Score 443.2; DB 6; Length 1662;
  Best Local Similarity 59.0%; Pred. No. 2.1e-66;
  Matches 798; Conservative   1; Mismatches 544; Indels   9; Gaps   2;

```
Qy        203 CTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAG 262

Db       1417 CTTAGATACGGGAGTTATGAGCGGAGCTCAGATATTTATAAGAGATGATCTAAAAATAAA 1358

Qy        263 TGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTT 322
```

```
Db       1357 CGACACTCAGATTGAAGTTTTGGCCGGAATCTTGAATCTTTGTGCACTCGTCGGATCACT 1298
Qy        323 CGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGT 382
Db       1297 AACGGCGGGGAAAACGTCTGACGTCATCGGCCGACGTTACACCATAGCTCTCTCCGCCGT 1238
Qy        383 CATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTT 442
Db       1237 GATATTCTTAGTGGGCTCAGTTCTTATGGGTTACGGCCCAAACTATCCTGTTTTGATGGT 1178
Qy        443 CGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACAC 502
Db       1177 CGGACGATGTATTGCCGGAGTTGGTGTAGGGTTTGCTCTCATGATAGCTCCGGTTTACTC 1118
Qy        503 CGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCAT 562
Db       1117 TGCCGAGATATCCTCTGCGTCACATAGAGGCTTTCTCACTTCTCTACCTGAGCTTTGTAT 1058
Qy        563 CAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAA 622
Db       1057 TAGTCTTGGGATTTTACTAGGCTATGTCTCGAATTACTGTTTTGGGAAACTGACGTTGAA 998
Qy        623 CCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCAT 682
Db        997 ACTTGGATGGAGATTGATGCTTGGAATCGCAGCGTTCCCGTCTTTGATATTGGCTTTTGG 938
Qy        683 GGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGC 742
Db        937 GATTACAAGAATGCCCGAGTCGCCGAGGTGGCTTGTGATGCAAGGTAGGTTAGAAGAAGC 878
Qy        743 CAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGA 802
Db        877 CAAGAAAATAATGGTTTTGGTATCCAATACGGAAGAAGAAGCTGAAGAACGGTTCAGAGA 818
Qy        803 CATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAA 862
Db        817 CATCTTAACCGCTGCGGAAGTAGATGTGACAGAGATAAAAGAAGTTGGTGGCGGCGTTAA 758
Qy        863 GAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGC 922
Db        757 GA---AGAAGAATCACGGTAAAAGTGTTTGGAGGGAGTTGGTGATAAAACCGCGACCTGC 701
Qy        923 CATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCAT 982
Db        700 GGTGCGTTTGATCTTGATAGCGGCCGTTGGGATACACTTTTTCGAGCATGCGACGGGGAT 641
Qy        983 TCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAA 1042
Db        640 TGAGGCCGTCGTGTTGTACAGCCCGAGAATTTTCAAGAAAGCCGGAGTTGTTTCGAAAGA 581
Qy       1043 ACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGC 1102
Db        580 CAAGCTCTTACTAGCCACGGTCGGTGTAGGTTTAACTAAGGCCTTTTTCATAATAATAGC 521
Qy       1103 GACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGAT 1162
Db        520 CACTTTCTTGCTCGACAAGGTAGGTCGTAGAAAGCTCCTGTTGACCAGTACGGGTGGGAT 461
Qy       1163 AATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGC 1222
Db        460 GGTCTTCGCATTGACCAGTTTAGCGGTTAGTCTCACGATGGT------TCAGCGATTTGG 407
Qy       1223 CAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTT 1282
Db        406 GCGGTTAGCGTGGGCATTGAGTTTAAGCATCGTCTCTACGTATGCATTTGTAGCTTTTTT 347
Qy       1283 CTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGT 1342
Db        346 CTCCATAGGGCTTGGACCTATCACATGGGTGTACAGCTCCGAGATATTCCCGTTGAGACT 287
Qy       1343 GCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTC 1402
Db        286 TAGGGCTCAGGGAGCGAGTATAGGTGTTGCGGTTAATAGGATCATGAACGCGACGGTCTC 227
```

```
Qy       1403 CATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTC 1462
Db        226 AATGAGTTTTCTGTCGATGACGAAGGCGATCACGACAGGGGGCGTGTTCTTCGTATTCGC 167
Qy       1463 CGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCG 1522
Db        166 CGGAATAGCGGTGGCGGCATGGTGGTTCTTCTTCTTTATGTTGCCGGAGACGAAAGGATT 107
Qy       1523 GACGCTGGAGGAGATGAGCAAGCTGTTCGGCG 1554
Db        106 GCCATTGGAGGAGATGGAGAAGCTTTTTGGTG 75
```

RESULT 15
ABZ12990
ID ABZ12990 standard; DNA; 1482 BP.
XX
AC ABZ12990;
XX
DT 21-JAN-2003 (first entry)
XX
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 795.
XX
KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX
OS Arabidopsis thaliana.
XX
PN WO200216655-A2.
XX
PD 28-FEB-2002.
XX
PF 24-AUG-2001; 2001WO-US026685.
XX
PR 24-AUG-2000; 2000US-0227866P.
PR 26-JAN-2001; 2001US-0264647P.
PR 22-JUN-2001; 2001US-0300111P.
XX
PA (SCRI) SCRIPPS RES INST.
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Harper JF, Kreps J, Wang X, Zhu T;
XX
DR WPI; 2002-304127/34.
XX
PT Identifying a stress condition to which a plant cell has been exposed and
PT producing plants with increased tolerance to these abiotic stresses.
XX
PS Claim 144; SEQ ID NO 795; 577pp + Sequence Listing; English.
XX
CC The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising: (a) contacting nucleic acid
CC representative of expressed polynucleotides in the plant cell with an
CC array or probes representative of the plant cell genome; and (b)
CC detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants
CC with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
CC in methods of the invention. Note: The sequence data for this patent is
CC not represented in the printed specification but is based on sequence
CC information supplied to Derwent by the European Patent Office
XX
SQ Sequence 1482 BP; 353 A; 311 C; 391 G; 427 T; 0 U; 0 Other;

Query Match 21.5%; Score 433.8; DB 6; Length 1482;
Best Local Similarity 57.8%; Pred. No. 8.4e-65;
Matches 809; Conservative 1; Mismatches 583; Indels 6; Gaps 2;

```
Qy        151 GGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATA 210
Db         44 GATTCGCTCTTCAATGTGCTATCGTCGCCTCCATCGTCTCCATCATCTTTGGTTACGATA 103
```

```
Qy       211 TCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGA 270
Db       104 CTGGTGTTATGAGTGGAGCGATGGTGTTTATAGAAGAAGATTTGAAGACAAACGACGTTC 163

Qy       271 AGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGG 330
Db       164 AAATAGAAGTTCTCACTGGAATTCTCAACCTTTGTGCCCTTGTCGGATCATTGCTCGCCG 223

Qy       331 GGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCT 390
Db       224 GAAGAACGTCGGACATAATCGGACGGCGTTACACAATCGTCTTGGCCTCAATACTATTCA 283

Qy       391 TCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCT 450
Db       284 TGTTAGGCTCAATATTGATGGGTTGGGGTCCGAATTATCCGGTTCTCCTATCCGGTAGAT 343

Qy       451 TCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGG 510
Db       344 GCACCGCTGGACTCGGAGTCGGTTTTGCTCTGATGGTTGCTCCGGTTTACTCTGCCGAGA 403

Qy       511 TGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCG 570
Db       404 TCGCAACTGCTTCACATAGAGGACTCTTAGCTTCTCTTCCTCACCTTTGTATCAGTATAG 463

Qy       571 GCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGT 630
Db       464 GGATTTTACTAGGTTATATCGTGAATTACTTCTTCTCCAAGTTACCTATGCATATCGGTT 523

Qy       631 GGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCG 690
Db       524 GGAGACTCATGCTCGGTATAGCCGCGGTTCCGTCCCTAGTGCTAGCGTTCGGGATCTTGA 583

Qy       691 GCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGG 750
Db       584 AAATGCCGGAATCTCCACGGTGGTTGATTATGCAAGGCCGTCTTAAGGAAGGCAAGGAGA 643

Qy       751 TGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGG 810
Db       644 TATTGGAATTGGTATCTAATTCCCCTGAAGAAGCAGAACTCCGCTTTCAAGACATCAAAG 703

Qy       811 CCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGA 870
Db       704 CTGCTGCGGGAATCGACCCGAAATGCGTAGACGATGTTGTGAAAATGG---AGGGTAAGA 760

Qy       871 GCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGC 930
Db       761 AGACTCATGGTGAAGGAGTGTGGAAAGAGCTCATTCTAAGACCAACTCCTGCAGTGAGAC 820

Qy       931 GCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCG 990
Db       821 GTGTTCTTTTAACTGCTCTTGGGATTCATTTCTTCCAACACGCCTCCGGAATCGAAGCAG 880

Qy       991 TCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCT 1050
Db       881 TGCTTTTATACGGTCCGAGGATCTTTAAGAAAGCAGGAATCACGACTAAAGACAAGCTTT 940

Qy      1051 TGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCT 1110
Db       941 TCTTGGTTACTATCGGTGTCGGAATCATGAAAACGACGTTTATTTTCACTGCGACTTTAT 1000

Qy      1111 TCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCT 1170
Db      1001 TGTTAGACAAGGTAGGTCGAAGGAAGCTTTTGTTGACCAGCGTTGGAGGAATGGTCATTG 1060

Qy      1171 CCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATAC 1230
Db      1061 CGTTGACAATGTTGGGATTTGGGCTTAC---AATGGCCCAAAATGCTGGCGGGAAATTAG 1117

Qy      1231 CTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCG 1290
Db      1118 CGTGGGCTTTAGTACTGAGCATAGTTGCGGCTTATAGTTTCGTGGCGTTTTTCTCTATTG 1177

Qy      1291 GCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGC 1350

Db      1178 GGCTCGGCCCAATAACTTGGGTCTACAGTTCTGAGGTTTTCCCGTTGAAGCTTAGGGCAC 1237

Qy      1351 TGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCT 1410
Db      1238 AAGGAGCGAGTCTCGGCGTTGCGGTGAACAGAGTAATGAACGCCACCGTGTCGATGTCGT 1297

Qy      1411 TCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCG 1470
Db      1298 TTTTGTCGTTGACTAGTGCGATAACCACCGGTGGAGCTTTCTTTATGTTCGCCGGAGTTG 1357

Qy      1471 CCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGG 1530
Db      1358 CGGCAGTGGCGTGGAATTTCTTCTTCTTCCTCTTGCCGGAGACGAAAGGAAAATCACTTG 1417

Qy      1531 AGGAGATGAGCAAGCTGTT 1549
Db      1418 AAGAAATCGAAGCGCTTTT 1436
```

Search completed: July 2, 2004, 02:47:01
Job time : 1163 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: July 2, 2004, 02:27:34 ; Search time 173 Seconds
(without alignments)
6470.150 Million cell updates/sec

Title: US-10-051-902A-21
Perfect score: 2017
Sequence: 1 cttacatgtaagctcgtgcc.....aaaaaaaaaaaaaaaaaaa 2017

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2016.6	100.0	2017	4	US-09-291-922-21	Sequence 21, Appl
2	1094.6	54.3	2089	4	US-09-291-922-25	Sequence 25, Appl
3	1005	49.8	1914	4	US-09-291-922-19	Sequence 19, Appl
4	981.6	48.7	1872	4	US-09-291-922-27	Sequence 27, Appl
5	575.2	28.5	1853	4	US-09-291-922-23	Sequence 23, Appl
6	343.6	17.0	615	4	US-09-291-922-17	Sequence 17, Appl
7	188.8	9.4	1675	4	US-09-679-686B-1	Sequence 1, Appli
8	168.4	8.3	1776	4	US-09-679-686B-11	Sequence 11, Appl
9	147.2	7.3	4411529	3	US-09-103-840A-1	Sequence 1, Appli
10	145.4	7.2	1545	4	US-09-489-039A-4731	Sequence 4731, Ap
11	141.2	7.0	1752	4	US-09-679-686B-17	Sequence 17, Appl
12	136	6.7	1167	4	US-09-679-686B-9	Sequence 9, Appli
13	128	6.3	4403765	3	US-09-103-840A-2	Sequence 2, Appli
14	118	5.9	1431	4	US-09-489-039A-4762	Sequence 4762, Ap
15	102	5.1	1344	4	US-09-252-991A-3992	Sequence 3992, Ap
16	102	5.1	1479	4	US-09-252-991A-3960	Sequence 3960, Ap
17	96.6	4.8	1689	4	US-10-162-012-45	Sequence 45, Appl
18	96.6	4.8	4385	4	US-10-162-012-43	Sequence 43, Appl
19	91.2	4.5	1320	4	US-09-252-991A-10535	Sequence 10535, A
20	91.2	4.5	1473	4	US-09-252-991A-10639	Sequence 10639, A
21	89.4	4.4	1407	4	US-09-252-991A-11395	Sequence 11395, A
22	89.4	4.4	2871	4	US-09-252-991A-11431	Sequence 11431, A
23	88.4	4.4	2601	4	US-09-291-922-7	Sequence 7, Appli
24	87.4	4.3	2824	4	US-09-291-922-1	Sequence 1, Appli
25	87.4	4.3	2856	4	US-09-643-597-135	Sequence 135, App
26	87.4	4.3	2856	4	US-09-480-884A-135	Sequence 135, App
27	87.4	4.3	2856	4	US-09-542-615A-135	Sequence 135, App
28	87.4	4.3	2856	4	US-09-606-421B-135	Sequence 135, App
29	87.4	4.3	2856	4	US-09-221-107-135	Sequence 135, App
c 30	87.2	4.3	534	4	US-09-252-991A-3916	Sequence 3916, Ap
31	86	4.3	1566	4	US-09-489-039A-2378	Sequence 2378, Ap
c 32	81.4	4.0	2058	4	US-09-252-991A-11442	Sequence 11442, A
33	80.8	4.0	1009	4	US-09-291-922-15	Sequence 15, Appl
34	78	3.9	1926	4	US-09-249-585A-4	Sequence 4, Appli
35	78	3.9	1931	2	US-09-130-114-2	Sequence 2, Appli
c 36	72.8	3.6	1002	4	US-09-252-991A-1099	Sequence 1099, Ap
37	72.8	3.6	1521	4	US-09-252-991A-989	Sequence 989, App
c 38	70.2	3.5	4403765	3	US-09-103-840A-2	Sequence 2, Appli
c 39	70.2	3.5	4411529	3	US-09-103-840A-1	Sequence 1, Appli
c 40	69.4	3.4	390	3	US-09-197-649-7	Sequence 7, Appli
41	69.2	3.4	1960	4	US-09-679-686B-15	Sequence 15, Appl
42	66.4	3.3	1288	1	US-08-440-856A-9	Sequence 9, Appli
c 43	65.8	3.3	603	4	US-09-252-991A-1063	Sequence 1063, Ap
44	65.8	3.3	822	4	US-09-252-991A-15614	Sequence 15614, A
45	65.4	3.2	1230	4	US-09-252-991A-947	Sequence 947, App

ALIGNMENTS

RESULT 1
US-09-291-922-21
; Sequence 21, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 21
; LENGTH: 2017
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-291-922-21

Query Match 100.0%; Score 2016.6; DB 4; Length 2017;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2017; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CTTACATGTAAGCTCGTGCCGGCACGAGCTTACACTCGACCGCCACTACTGTACACGGCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTTACATGTAAGCTCGTGCCGGCACGAGCTTACACTCGACCGCCACTACTGTACACGGCC 60

Qy         61 CAGAGCGAGCCTCCTCCTCCTCTGCACCACCGGAGATGGCTTCCGCCGCGCTGCCGGAGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGAGCGAGCCTCCTCCTCCTCTGCACCACCGGAGATGGCTTCCGCCGCGCTGCCGGAGG 120

Qy        121 CCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCT 180

Qy        181 CCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACA 240

Qy        241 TCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACC 300

Qy        301 TCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGT 360
```

```
Qy        361 ACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGGSGTTCCTCATGGGGTTCGCCG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGGSGTTCCTCATGGGGTTCGCCG 420

Qy        421 TCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGC 480

Qy        481 TCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGA 540

Qy        541 CGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATG 600

Qy        601 CTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGC 660

Qy        661 CGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCA 720

Qy        721 TGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGG 780

Qy        781 AGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACG 840

Qy        841 GCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGC 900

Qy        901 TCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACT 960

Qy        961 TCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGA 1020

Qy       1021 GCCCCGGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GCCCCGGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCA 1080

Qy       1081 AGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGT 1140

Qy       1141 TGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCG 1200

Qy       1201 TCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCA 1260

Qy       1261 CCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCT 1320

Qy       1321 CGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACC 1380

Qy       1381 GCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 GCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCG 1440

Qy       1441 GCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCT 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCT 1500

Qy       1501 ACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 ACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGG 1560

Qy       1561 CCGCCGCCTCGGAATCAGACGAGCCAGCCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCA 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CCGCCGCCTCGGAATCAGACGAGCCAGCCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCA 1620

Qy       1621 CTAACTGATCAAACTAACCGCAAAATCACCAAATCCTAAGGGTTTTCTTGCAAAAACGTG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 CTAACTGATCAAACTAACCGCAAAATCACCAAATCCTAAGGGTTTTCTTGCAAAAACGTG 1680

Qy       1681 TGCTGTACTGGCTAGCTAGCAAGTAGTAGCAGCAACGTGGGAAGATTCGCTGATCCGGCG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 TGCTGTACTGGCTAGCTAGCAAGTAGTAGCAGCAACGTGGGAAGATTCGCTGATCCGGCG 1740

Qy       1741 TTGCTGGAGAGCGACGGCCGGCGACGACAAAGCTGAGCTCCAGCTCGAGACTTCTTAAAA 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TTGCTGGAGAGCGACGGCCGGCGACGACAAAGCTGAGCTCCAGCTCGAGACTTCTTAAAA 1800

Qy       1801 TCATCTTCAAGTACATGGATTTTATTTTGCTCTTTGCTTTGTCCGTAAAAGTTGTACTAT 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 TCATCTTCAAGTACATGGATTTTATTTTGCTCTTTGCTTTGTCCGTAAAAGTTGTACTAT 1860

Qy       1861 GCGATGAAGAATACCAGTATGTAGCAAGGCTGAGGTTGTGTGTAGCTACTAGAAGTGTCA 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 GCGATGAAGAATACCAGTATGTAGCAAGGCTGAGGTTGTGTGTAGCTACTAGAAGTGTCA 1920

Qy       1921 GTCACGTTGTTCTTGTAAGAAATGTTTAACTGTTAATTAAGCAGTATTGTTGCAGTAATC 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GTCACGTTGTTCTTGTAAGAAATGTTTAACTGTTAATTAAGCAGTATTGTTGCAGTAATC 1980

Qy       1981 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2017
              |||||||||||||||||||||||||||||||||||||||||||
Db       1981 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2017
```

RESULT 2
US-09-291-922-25
; Sequence 25, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 25
; LENGTH: 2089
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-291-922-25

Query Match 54.3%; Score 1094.6; DB 4; Length 2089;
Best Local Similarity 79.1%; Pred. No. 8.7e-206;
Matches 1343; Conservative 1; Mismatches 335; Indels 18; Gaps 3;

```
Qy         84 GCACCACCGGAGATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGC 143
              ||  |  ||  |||||    |||||| || |||| ||||||||| ||| ||||||||||||
Db         65 GCGTCGTCGACGATGGACCGCGCCGCACTCCCGGCGGCCGTCGAGCCCAAGAAGAAGGGC 124

Qy        144 AACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGC 203
```

```
Db        125 AACGTGAGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGC 184
Qy        204 TACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGT 263
Db        185 TACGACATCGGCGTGATGAGCGGAGCGTCGCTGTACATCCAGAAGGATCTGAAGATCAAC 244
Qy        264 GACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTC 323
Db        245 GACACCCAGCTGGAGGTCCTCATGGGCATCCTCAACGTGTACTCGCTCATTGGCTCCTTC 304
Qy        324 GCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTC 383
Db        305 GCGGCGGGGCGGACGTCCGACTGGATCGGCCGGCGCTTCACCATCGTCTTCGCCGCCGTC 364
Qy        384 ATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTC 443
Db        365 ATCTTCTTCGCGGGCGCCCTCATCATGGGCTTCTCCGTCAACTACGCCATGCTCATGTTC 424
Qy        444 GGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACC 503
Db        425 GGGCGCTTCGTGGCCGGCATCGGCGTGGGGTACGCTCTCATGATCGCGCCCGTGAACACG 484
Qy        504 GCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATC 563
Db        485 GGCGAGGTGTCCCCCGCGTCTGCCCGTGGGGTTCTCACATCCTTCCCGGAGGTGTTCATC 544
Qy        564 AACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAAC 623
Db        545 AACTTCGGCATCCTCCTCGGATATGTCTCCAACTTCGCCTTCGCCCGCCTCTCCCTCCGC 604
Qy        624 CTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATG 683
Db        605 CTCGGCTGGCGCATTATGCTCGGCATAGGCGCGGTGCCGTCCGTCCTGCTCGCGTTCATG 664
Qy        684 GTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCC 743
Db        665 GTGCTCGGCATGCCCGAGTCTCCCCGGTGGCTCGTCATGAAGGGCCGTCTCGCGGACGCC 724
Qy        744 AAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGAC 803
Db        725 AAGGTTGTGCTTGCCAAGACGTCCGACACGCCGGAAGAGGCCGCCGAGCGCATCGCCGAC 784
Qy        804 ATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCC--- 860
Db        785 ATTAAGACTGCCGCCGGCATCCCTCTGGGCCTCGACGGCGACGTGGTCCCCGTGCCCAAA 844
Qy        861 AAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCG 920
Db        845 AACAAAGGAAGCAGCGAGGAGAAGCGCGTTTTGAAGGACCTCATCCTGTCACCGACCATA 904
Qy        921 GCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGC 980
Db        905 GCCATGCGCCACATCCTCATCGCGGGAATCGGCATCCACTTCTTCCAGCAGTCTTCGGGC 964
Qy        981 ATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGAC 1040
Db        965 ATCGACGCCGTCGTGCTCTACAGCCCGCTAGTTTTCAAGAGCGCCGGCATCACGGGCGAC 1024
Qy       1041 AAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTG 1100
Db       1025 AGCCGTCTCCGCGGCACCACCGTGGCGGTCGGGGCCACCAATACGGTCTTCATCCTGGTG 1084
Qy       1101 GCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGG 1160
Db       1085 GCCACCTTCCTCCTCGACCGCATCCGCCGGCGGCCGCTGGTGCTGACCAGCACGGGCGGC 1144
Qy       1161 ATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGAC 1220
Db       1145 ATGCTCGTCTCCTTAGTGGGCCTCGCGACGGGGCTCACCGTCATCAGCCGCCACCCGGAC 1204
Qy       1221 GCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTC 1280
```

```
Db        1205 GAGAAGATCACCTGGGCCATCGTCCTGTGCATCTTCTGCATCATGGCCTACGTGGCCTTC 1264
Qy        1281 TTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAG 1340
Db        1265 TTCTCCATCGGCCTCGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTGCAC 1324
Qy        1341 GTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATC 1400
Db        1325 GTGCGCGCGCTGGGCTGCTCCCTGGGCGTGGCCGTCAACCGCCTGACCAGCGGCGTGATC 1384
Qy        1401 TCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTAC 1460
Db        1385 TCCATGACCTTCATTTCGCTGTCCAAGGCCATGACCATCGGCGGCGCCTTCTTCCTCTTC 1444
Qy        1461 TCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGC 1520
Db        1445 GCCGGCATCGCCTCATTCGCATGGGTGTTCTTCTTCGCCTACCTGCCGGAGACCCGCGGC 1504
Qy        1521 CGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGAC 1580
Db        1505 CGCACGCTGGAGGACATGAGCTCGCTGTTCGGCAACACGGCCACGCACAAGCAGGGCGCC 1564
Qy        1581 GAGCCAGCCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCACTAACTGATCAAACTAACCG 1640
Db        1565 GCGGAAGCCGACGACGACGCCGGGGAGAAGAAGGTGGAAATGGCCGCCACCAACTGACCG 1624
Qy        1641 CAAAATCACCAAATCCTAAGGGTTTTCTTGCAAAAACGTGTGCTGTACTGGCTAGCTAGC 1700
Db        1625 CAAGTTGGCAGATCGCGATGCGAAGACTTGCGCTGTATCCGTCTCGGCTAGCTAGCTGCC 1684
Qy        1701 A-----------AGTAGTAGCAGCAACGTGGGAAGATTCGCTGATCCGGC----GTTGCT 1745
Db        1685 ACAAGGCCACATAGATGACGAAGTAGCGTGGGAAGATTCGCTGATCCGGCCGGAGCTGCC 1744
Qy        1746 GGAGAGCGACGGCCGGC 1762
Db        1745 GGAGGGCGACGGCAAGC 1761
```

RESULT 3
US-09-291-922-19
; Sequence 19, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 19
; LENGTH: 1914
; TYPE: DNA
; ORGANISM: Zea mays
US-09-291-922-19

Query Match 49.8%; Score 1005; DB 4; Length 1914;
Best Local Similarity 79.2%; Pred. No. 3.2e-188;
Matches 1220; Conservative 1; Mismatches 311; Indels 9; Gaps 2;

```
Qy          95 GATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTT 154
Db          55 GATGGCTTCCGACGAGCTCGCAAAGGCCGTCGAGCCCAGGAAGAAGGGCAACGTCAAGTA 114
Qy         155 CGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGG 214
Db         115 TGCCTCCATATGTGCCATCCTGGCCTCCATGGCCTCTGTCATCCTTGGCTATGACATTGG 174
```

```
Qy       215 GGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGT 274
Db       175 GGTGATGAGTGGAGCGGCCATGTACATCAAGAAGGACCTGAATATCACGGACGTGCAGCT 234

Qy       275 GGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCG 334
Db       235 GGAGATCCTGATCGGGATCCTCAGTCTCTACTCGCTGTTCGGATCCTTCGCTGGCGCGCG 294

Qy       335 GACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGC 394
Db       295 GACGTCCGACAGGATCGGGCGCCGCTTGACCGTCGTGTTCGCCGCTGTCATCTTCTTCGT 354

Qy       395 GGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGT 454
Db       355 GGGCTCGTTGCTCATGGGTTTCGCCGTCAACTACGGCATGCTCATGGCGGGCCGCTTCGT 414

Qy       455 GGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTC 514
Db       415 GGCCGGAGTCGGTGTGGGCTACGGGGGCATGATCGCGCCCGTGTACACGGCCGAGATCTC 474

Qy       515 GCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCAT 574
Db       475 GCCTGCGGCGTCCCGTGGCTTCCTGACCACCTTCCCGGAGGTGTTCATCAACATCGGCAT 534

Qy       575 CCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCG 634
Db       535 CCTGCTTGGCTACCTGTCCAACTTCGCGTTCGCGCGCCTCCCGCTCCACCTCGGCTGGCG 594

Qy       635 CATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCAT 694
Db       595 CGTCATGCTCGCCATTGGCGCAGTTCCGTCCGGCCTGCTCGCGCTCCTGGTGTTCTGCAT 654

Qy       695 GCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCT 754
Db       655 GCCCGAGTCGCCTCGGTGGCTGGTCTTGAAGGGCCGCCTCGCGGACGCCAGGGCTGTGCT 714

Qy       755 GGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGC 814
Db       715 AGAGAAGACCTCTGCCACGCCAGAGGAGGCCGCCGAGCGGCTGGCCGACATCAAGGCCGC 774

Qy       815 CGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCC---CAAGAGAGGGAG 871
Db       775 GGCGGGGATTCCGAAGGGCCTCGACGGGGACGTAGTCACCGTACCCGGCAAGGAGCAAGG 834

Qy       872 CGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCG 931
Db       835 CGGCGGTGAGTTGCAGGTGTGGAAGAAGCTCATCCTGTCCCCGACCCCGGCTGTCCGACG 894

Qy       932 CATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGT 991
Db       895 CATACTGCTCTCGGCCGTGGGTCTCCACTTCTTCCAGCAGGCTTCTGGCAGCGACTCCGT 954

Qy       992 CGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTT 1051
Db       955 CGTCCAGTACAGCGCCCGCCTGTTCAAGAGCGCGGGGATCACCGACGACAACAAGCTCCT 1014

Qy      1052 GGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTT 1111
Db      1015 GGGCGTCACCTGCGCGGTGGGCGTGACCAAGACGTTCTTCATCCTGGTGGCCACGTTCCT 1074

Qy      1112 CATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTC 1171
Db      1075 GCTGGACCGCGCGGGGCGTCGGCCTCTGCTGCTGATCAGCACGGGCGGGATGATTGTCTC 1134

Qy      1172 CCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACC 1231
Db      1135 GCTCATCTGCCTCGGGTCGGGGCTCACCGTCGCGGGGCATCACCCGGACACCAAGGTCGC 1194

Qy      1232 TTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGG 1291
Db      1195 GTGGGCCGTCGCCCTGTGCATCGCGTCAACCCTGTCCTACATCGCCTTCTTCTCCATCGG 1254

Qy      1292 CCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCT 1351
Db      1255 CCTCGGGCCCATCACGGGCGTGTACACCTCGGAAATATTCCCGCTGCAGGTGCGCGCGCT 1314

Qy      1352 GGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTT 1411
Db      1315 GGGCTTCGCGGTGGGTGTGGCGAGCAACCGCGTCACCAGCGCCGTCATCTCCATGACCTT 1374

Qy      1412 CCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGC 1471
Db      1375 CCTGTCCCTCTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGC 1434

Qy      1472 CGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGA 1531
Db      1435 CGCGGTCGCTTGGGTTTTCTTCTTCACGTGCCTCCCGGAGACACGCGGCCGGACGCTGGA 1494

Qy      1532 GGAGATGAGCAAGCTGTTCGGC------GACACGGCCGCCGCCTCGGAATCAGACGAGCC 1585
Db      1495 GGAGATGGGCAAGCTGTTCGGCATGCCAGACACGGGCATGGCTGAAGAAGCAGAAGACGC 1554

Qy      1586 AGCCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCACTAACT 1626
Db      1555 CGCAGCCAAGGAGAAGGTGGTGGAACTGCCTAGCAGCAAGT 1595
```

RESULT 4
US-09-291-922-27
; Sequence 27, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 27
; LENGTH: 1872
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-291-922-27

Query Match 48.7%; Score 981.6; DB 4; Length 1872;
Best Local Similarity 77.4%; Pred. No. 1.2e-183;
Matches 1188; Conservative 1; Mismatches 345; Indels 0; Gaps 0;

```
Qy        99 GCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCC 158
Db       128 GCTGCGCTCCCGGAGCCGGGGGCAGTCCATCCAAGGAACAAGGGCAATTTCAAGTACGCC 187

Qy       159 TTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTG 218
Db       188 TTCACCTGCGCCCTCTGTGCTTCCATGGCCACCATCGTCCTCGGCTACGACGTTGGGGTG 247

Qy       219 ATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAG 278
Db       248 ATGAGCGGTGCGTCGCTGTACATCAAGAGGGACCTGCAGATCACGGACGTGCAGCTGGAG 307

Qy       279 GTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACG 338
Db       308 ATCATGATGGGCATCCTGAGCGTGTACGCGCTCATCGGGTCCTTCCTCGGCGCGAGGACG 367

Qy       339 TCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGG 398
Db       368 TCCGACTGGGTCGGCCGCCGCGTCACCGTCGTCTTCGCGGCCGCCATCTTCAACAACGGC 427

Qy       399 GSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCC 458
```

```
Db        428 TCCTTGCTCATGGGCTTCGCGGTCAACTACGCCATGCTCATGGTCGGGCGCTTCGTCACC 487

Qy        459 GGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCG 518

Db        488 GGAATCGGCGTGGGCTACGCCATCATGGTCGCGCCAGTGTACACGCCCGAGGTGTCCCCG 547

Qy        519 GCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTG 578

Db        548 GCGTCGGCCCGCGGCTTCCTCACGTCTTTCACCGAGGTGTTCATCAATGTGGGCATCCTC 607

Qy        579 CTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATC 638

Db        608 CTTGGCTACGTCTCCAACTACGCCTTCGCGCGCCTCCCGCTCCACCTCAGCTGGCGCGTC 667

Qy        639 ATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCG 698

Db        668 ATGCTCGGCATCGGCGCCGTCCCGTCCGCCCTGCTTGCGCTCATGGTGTTCGGCATGCCG 727

Qy        699 GAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAG 758

Db        728 GAGTCTCCTCGCTGGCTCGTCATGAAAGGCCGCCTCGCGGACGCCAGGGCCGTTCTGGCC 787

Qy        759 AAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCC 818

Db        788 AAGACCTCCGACACGCCGGAGGAGGCCGTGGAGCGCCTTGACCAGATCAAGGCTGCCGCC 847

Qy        819 GGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAAC 878

Db        848 GGCATCCCTAGGGAACTTGACGGCGACGTGGTCGTCATGCCTAAGACAAAAGGCGGCCAG 907

Qy        879 GAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTG 938

Db        908 GAGAAGCAGGTGTGGAAGGAGCTCATCTTTTCGCCGACCCCAGCCATGCGGCGCATACTG 967

Qy        939 CTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTC 998

Db        968 CTCGCGGCGCTCGGCATCCATTTCTTTCAGCAGGCGACGGGCTCCGACTCCGTCGTGCTC 1027

Qy        999 TACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGCACC 1058

Db       1028 TATAGCCCACGCGTGTTCCAGAGCGCGGGCATCACCGGCGACAACCACCTGCTCGGCGCC 1087

Qy       1059 ACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGAC 1118

Db       1088 ACATGCGCCATGGGGGTCATGAAGACGCTCTTCATCCTGGTGGCCACGTTCCAGCTCGAC 1147

Qy       1119 GGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATC 1178

Db       1148 CGCGTCGGCAGGCGGCCGCTGCTGCTGACCAGCACGGCCGGCATGCTCGCCTGTCTCATC 1207

Qy       1179 GGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCC 1238

Db       1208 GGCCTCGGGACGGGCCTCACCGTCGTGGGTCGGCACCCGGACGCCAAGGTCCCGTGGGCC 1267

Qy       1239 ATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGC 1298

Db       1268 ATCGGCCTGTGCATCGTGTCCATCTTGGCCTACGTGTCCTTCTTCTCCATCGGCCTCGGG 1327

Qy       1299 CCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGC 1358

Db       1328 CCCCTCACCAGCGTGTACACCTCGGAGGTCTTCCCACTGCGGGTGCGCGCGCTGGGCTTC 1387

Qy       1359 TCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCG 1418

Db       1388 GCGCTGGGCACGTCATGCAACCGCGTCACCAGCGCCGCGGTCTCCATGTCCTTCCTGTCC 1447

Qy       1419 CTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTC 1478

Db       1448 TTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTGTACGCCGGCATCGCGGCGATA 1507

Qy       1479 GCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATG 1538

Db       1508 GGATGGATTTTCTTCTTCACCTTCATTCCGGAGACGCGTGGCCTGCCGCTCGAGGAGATA 1567

Qy       1539 AGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAGCCAAGGAGAAG 1598

Db       1568 GGGAAGCTTTTCGGCATGACGGACACGGCCGTCGAAGCCCAAGACACCGCCACGAAAGAC 1627

Qy       1599 AAGAAGGTGGAAATGGCCGCCACTAACTGATCAA 1632

Db       1628 AAGGCGAAAGTAGGGGAGATGAACTAGTGAGCTA 1661
```

RESULT 5
US-09-291-922-23
; Sequence 23, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 23
; LENGTH: 1853
; TYPE: DNA
; ORGANISM: Glycine max
US-09-291-922-23

Query Match 28.5%; Score 575.2; DB 4; Length 1853;
Best Local Similarity 63.6%; Pred. No. 5.9e-104;
Matches 909; Conservative 2; Mismatches 510; Indels 9; Gaps 2;

```
Qy         109 CGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGCG 168

Db         210 CACTTCAGGATTTCGATCCTCCAAAGAAGCGCAAAAGGAACAAGTATGCTTTTGCTTGTG 269

Qy         169 CCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGGG 228

Db         270 CTATGCTGGCCTCCATGACTTCCATCTTGCTTGGTTATGATATTGGAGTGATGAGTGGAG 329

Qy         229 CGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGG 288

Db         330 CAGCCATATACATAAAAAGGGACCTGAAAGTCTCGGACGAGCAAATCGAGATCCTGCTCG 389

Qy         289 GCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGA 348

Db         390 GAATCATCAACCTATACTCTCTGATAGGCTCATGTCTCGCCGGCAGAACCTCCGACTGGA 449

Qy         349 TCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGSGTTCCTCA 408

Db         450 TAGGTCCCCGTTACACGATTGTTTTCGCCGGCACCATCTTCTTTGTCGGAGCACTTCTCA 509

Qy         409 TGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCG 468

Db         510 TGGGTTTCTCCCCCAATTATTCCTTTCTCATGTTTGGCCGTTTCGTCGCTGGCATTGGCA 569

Qy         469 TGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGC 528

Db         570 TCGGCTACGCCCTCATGATAGCCCCCGTCTACACCGCCGAGGTCTCCCCGGCCTCCTCTC 629

Qy         529 GTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACG 588

Db         630 GTGGCTTCCTCACTTCCTTCCCTGAGGTATTTATTAATGGAGGGATATTAATTGGATACA 689

Qy         589 TCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCA 648

Db         690 TATCAAACTATGCATTTTCGAAGCTGACACTAAAGGTGGGATGGCGAATGATGCTTGGAG 749
```

```
Qy        649 TCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGC 708
              | || ||    || || || || ||  |  |    ||| | |  |||||||||| ||
Db        750 TTGGTGCAATACCTTCGGTACTCCTAACAGTAGGAGTGTTGGCGATGCCGGAGTCCCCAA 809

Qy        709 GGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCG 768
              ||||||| || |||| ||| ||  | |  || || |     |||||  | || ||||| |
Db        810 GGTGGCTTGTGATGAGGGGTCGTTTGGGAGAGGCAAGAAAAGTGCTTAACAAAACCTCAG 869

Qy        769 ACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTG 828
              |||    ||| ||||||       | || || || |||||    |||||| || ||||| |
Db        870 ACAGCAAGGAAGAGGCCCAACTAAGGCTAGCGGAAATCAAACAAGCCGCAGGGATCCCCG 929

Qy        829 AGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGG 888
              ||      | ||| |||||| ||     ||    ||   | | | |||  | |      | |
Db        930 AGAGTTGCAACGACGACGTCGTTCAGGTAAATAAACAAAGCAACGGTGAAG------GTG 983

Qy        889 TGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGA 948
              | ||||| |||||| |||| |  || || || || || || | |||| |   |  | |
Db        984 TATGGAAAGAGCTCTTCCTCTATCCAACGCCCGCAATTCGTCACATCGTAATCGCTGCCC 1043

Qy        949 TCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTC 1008
              | || || |||||||||||| || |||| |||| |  || ||||||| || ||||||||
Db       1044 TTGGTATTCACTTCTTCCAACAAGCGTCGGGCGTAGACGCCGTCGTTTTGTACAGCCCCA 1103

Qy       1009 TCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCGT 1068
                | ||| | |   | ||  |  || |||||||  ||     |    | ||   | |
Db       1104 GGATCTTCGAAAAGGCTGGGATTACAAACGACACGCATAAGCTTCTTGCAACCGTGGCCG 1163

Qy       1069 TCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGC 1128
              | ||  ||   ||||   | ||||||||||  ||| || ||      | ||| |||| || |
Db       1164 TTGGATTCGTTAAGACCGTGTTCATCTTGGCGGCTACGTTTACGTTGGACCGCGTGGGTC 1223

Qy       1129 GGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCG 1188
              | || ||| |||| ||   ||     ||||| ||  | ||||| ||  ||   || |
Db       1224 GTCGTCCGTTGTTATTGTCTAGTGTCGGCGGCATGGTGCTCTCGCTTCTCACGCTTGCGA 1283

Qy       1189 CCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAA 1248
              | | ||||| ||  | |   | ||  | ||    || ||   |||||| | ||     |
Db       1284 TCAGCCTCACTGTTATTG---ATCATTCGGAGAGGAAATTAATGTGGGCCGTTGGATCGA 1340

Qy       1249 GCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGT 1308
              |||| |||     | || ||||| |||    ||||||||||     || ||||||||||
Db       1341 GCATAGCCATGGTGTTGGCTTACGTGGCCACGTTCTCCATCGGTGCGGGTCCCATCACGT 1400

Qy       1309 GGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCG 1368
              |||| || || || ||||||||||||  |  | |||| |||| :||    |    || |
Db       1401 GGGTCTATAGTTCTGAGATCTTCCCGTTGAGGCTGCGGGCGCARGGTGCGGCCGCGGGAG 1460

Qy       1369 TCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGG 1428
              | ||  |  ||  |  ||| |||| ||  |||| ||||| || ||||| ||  | |  |
Db       1461 TTGCGGTGAATAGGACCACTAGCGCGGTTGTCTCAATGACTTTTCTGTCCCTCACTAGAG 1520

Qy       1429 CCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGT 1488
              ||||||| || || ||    |||||||| || |  ||||| ||  |  | |  ||| | |
Db       1521 CCATCACTATTGGTGGAGCTTTCTTCCTTTATTGTGGCATTGCTACTGTTGGGTGGATAT 1580

Qy       1489 TCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATG 1538
              |||| ||||||  | | || |||||||| ||    ||||| || || |||
Db       1581 TCTTTTACACCGTCTTGCCTGAGACCCGGGGAAAAACGCTCGAAGACATG 1630
```

RESULT 6
US-09-291-922-17
; Sequence 17, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 17
; LENGTH: 615
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (149)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (271)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (304)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (334)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (357)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (476)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (599)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (602)
US-09-291-922-17

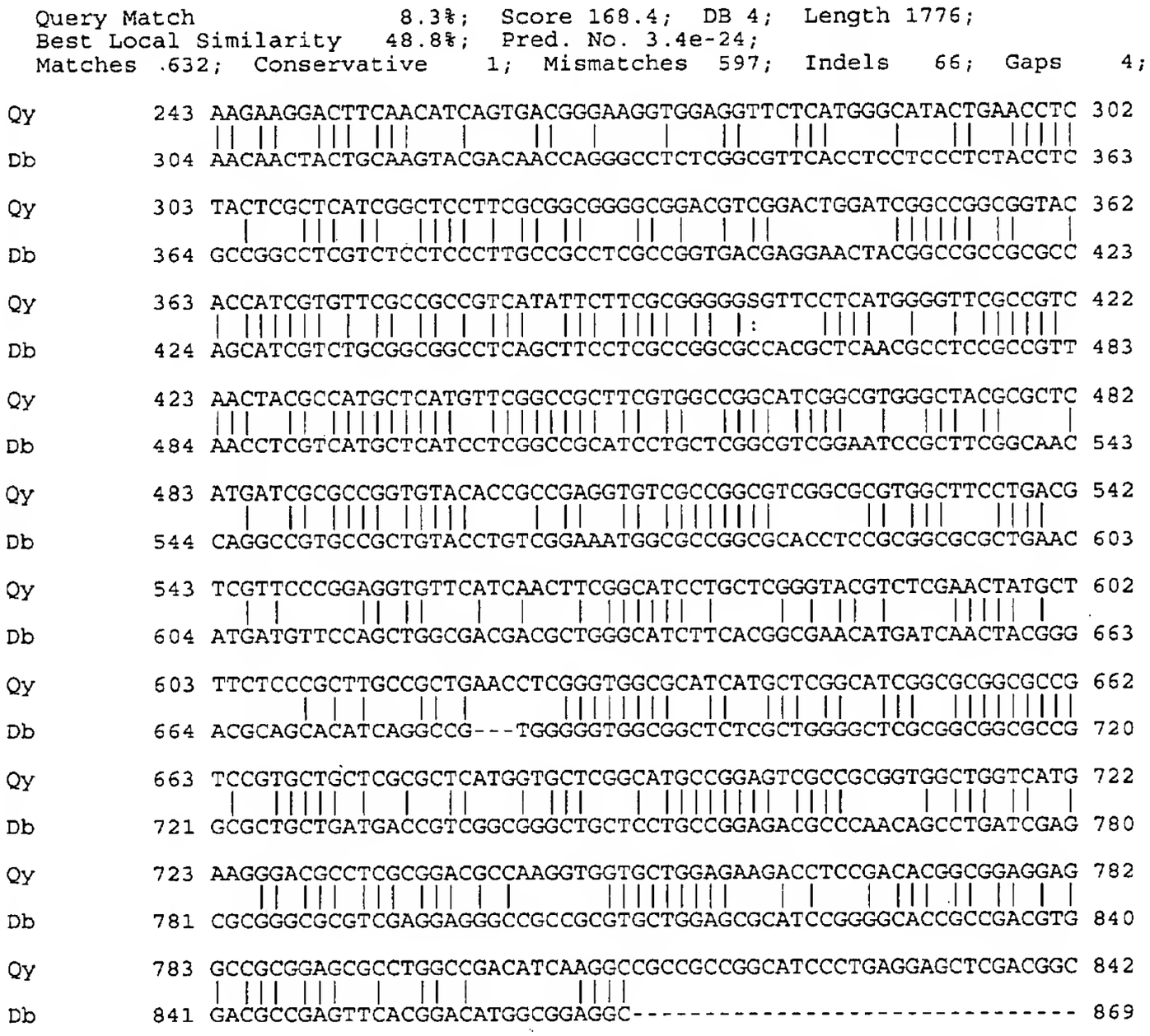
Query Match 17.0%; Score 343.6; DB 4; Length 615;
Best Local Similarity 85.6%; Pred. No. 1.1e-58;
Matches 423; Conservative 1; Mismatches 66; Indels 4; Gaps 4;

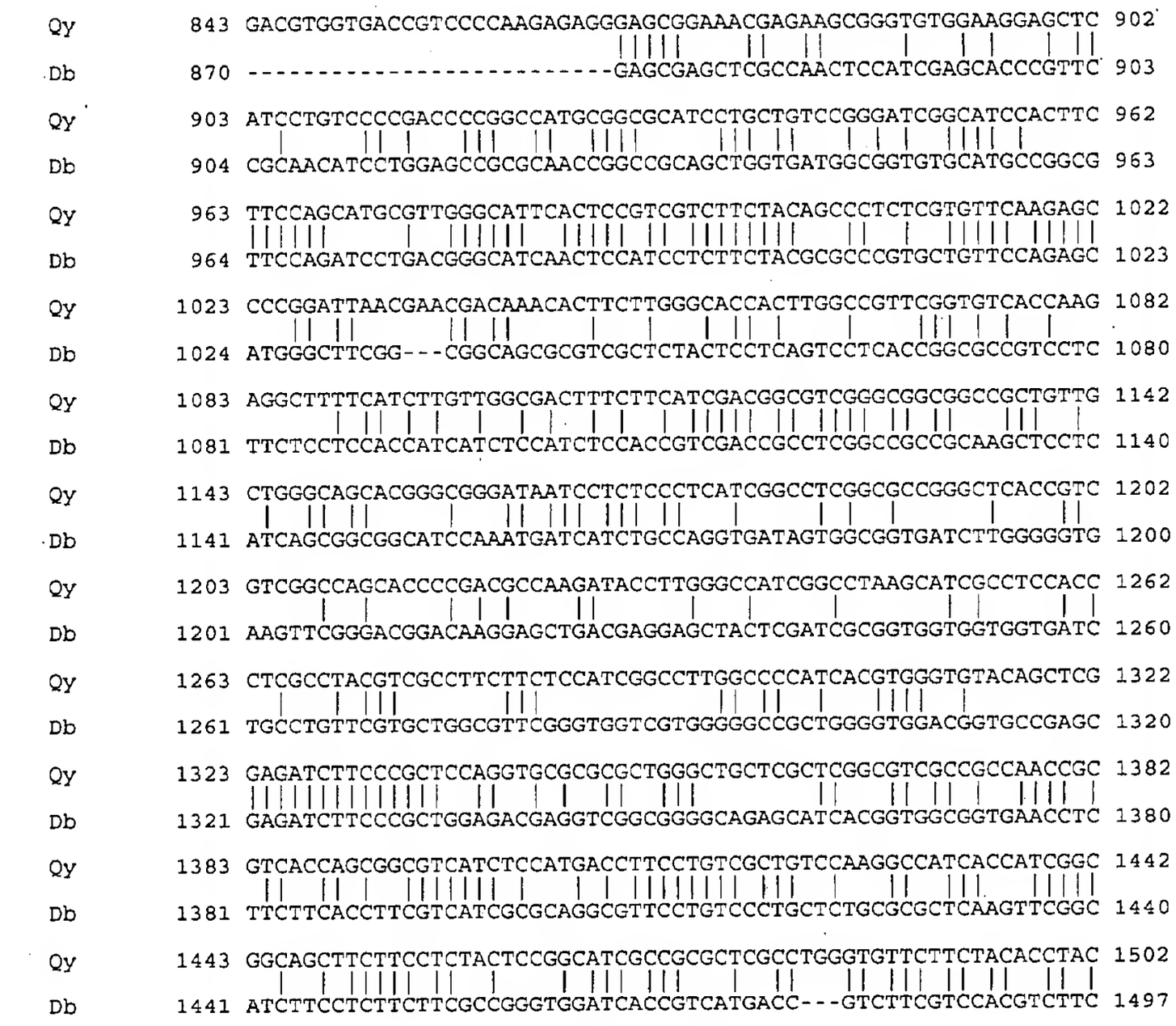
```
Qy         84 GCACCACCGGAGATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGC 143
              | |||  |  |||||||||||  |||||||| |||| ||| |||   |||||| |||
Db         65 GAACCGTCAACGATGGCTTCCGCTCCGCTGCCGGCGGCCATCGAGCCCGGGAAGAAAGGC 124

Qy        144 AACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGC 203
              ||||||  |||||||||||||||| ||||||||||| ||||||||||||| ||||||
Db        125 AACGTCAAGTTCGCCTTCGCCTGCNCCATCCTCGCCTCAATGACCTCCATCCTTCTCGGC 184

Qy        204 TACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGT 263
              || |||||||  |||||||||| |||||| |||||||||||||| | || |||||
Db        185 TATGATATCGGAGTGATGAGCGGCGCGTCGTTGTACATCAAGAAGGACCTGAAAATCAGC 244

Qy        264 GACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTC 323
              |||| |||| ||||| | || ||||| || || ||| | ||||||||||||||||| ||
Db        245 GACGTGAAGCTGGAGATCCTGATGGGNATCCTCAACGTGTACTCGCTCATCGGCTCGTTN 304

Qy        324 GCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTC 383
              ||||| |||||||||||  |||||||||||  | || |||||||||||||||||| || ||
Db        305 GCGGCAGGGCGGACGTCCGACTGGATCGGNC-GCCGTACACCATCGTGTTCGCNGCGGTG 363

Qy        384 ATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTC 443
              || ||||||||||  |: |||||||||| |||||||  ||||||    ||||||||||||
Db        364 ATCTTCTTCGCGGGCGC-TTCCTCATGGGCTTCGCCGTGAACTACTGGATGCTCATGTTC 422

Qy        444 GGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACC 503
              || |||||||||||||  |||||||||||||||||||||||||||||   | || |||||
Db        423 GGGCGCTTCGTGGCCGGGATCGGCGTGGGCTACGCGCTCATGATCGCAACCGTNTACACG 482

Qy        504 GCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATC 563
              ||||| |||||  || || |||||  || ||||||||||||||||||||||||||||||
Db        483 GCCGAAGTGTC-CCCGCATCGGCCCGCGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATC 541

Qy        564 AACTTCGGCATCCT 577
```

```
Db        542 -ACTTCGGCATCCT 554

RESULT 7
US-09-679-686B-1
; Sequence 1, Application US/09679686B
; Patent No. 6624343
; GENERAL INFORMATION:
;  APPLICANT: Allen, Stephen M.
;  APPLICANT:  Lightner, Jonathan E.
;  APPLICANT:  Rafalski, J. Antoni
;  APPLICANT:  Thorpe, Catherine J.
;  TITLE OF INVENTION: HEXOSE CARRIER PROTEINS
;  FILE REFERENCE: BB1160 US NA
;  CURRENT APPLICATION NUMBER: US/09/679,686B
;  CURRENT FILING DATE:  2003-01-16
;  PRIOR APPLICATION NUMBER: 60/081,131
;  PRIOR FILING DATE: 1998-04-09
;  PRIOR APPLICATION NUMBER: PCT/US99/07561
;  PRIOR FILING DATE: 1999-04-07
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
;   LENGTH: 1675
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1624)
;   OTHER INFORMATION: n = a, c, g or t
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1638)
;   OTHER INFORMATION: n = a, c, g or t
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1668)
;   OTHER INFORMATION: n = a, c, g or t
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1670)
;   OTHER INFORMATION: n = a, c, g or t
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1674)
;   OTHER INFORMATION: n = a, c, g or t
US-09-679-686B-1

  Query Match             9.4%;  Score 188.8;  DB 4;  Length 1675;
  Best Local Similarity   50.1%;  Pred. No. 3.3e-28;
  Matches  660;  Conservative    1;  Mismatches  588;  Indels   69;  Gaps    5;

Qy        227 GGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCAT 286
              ||||   ||| ||    | | |  ||| |||   |    ||  |  | ||  | |  |||
Db        368 GGCGCACCTGCACGAGACGGACTACTGCAAGTACGACAACCAGGTGCTGACGCTCTTCAC 427

Qy        287 GGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTG 346
                ||  ||  ||      ||| ||  | ||||||  ||          | |  |  |
Db        428 CTCGTCGCTCTACTTCGCGGGCCTCGTCTCCACCTTCGGCGCCTCCTACGTGACCAAGCG 487

Qy        347 GATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGSGTTCCT 406
              |||||||||   || ||||  ||  | |||||  ||  ||||||   ||  |:   | |
Db        488 CCACGGCCGGCGCGCCAGCATCATGGGTGGCGCCGCCAGCTTCTTCCTCGGCGGCGCCAT 547

Qy        407 CATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGG 466
              ||  ||   |||| | |||  ||||||||||||  |||| ||| || |   |||| ||||
Db        548 CAACGGCGCCGCCATGAACATCGCCATGCTCATCGTCGGACGCATCCTCCTCGGCGTCGG 607

Qy        467 CGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGC 526
              ||| |||| |||      |  || ||| ||||||    | ||| || ||||||||
Db        608 CGTCGGCTTCGCAAATCAGGCCGTGCCTGTGTACCTGTCGGAGATGGCGCCGGCGCGTCT 667

Qy        527 GCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTA 586
              || ||| | || |       |    || || | |||| | ||||||||||   | | |
Db        668 CCGGGGCATGCTCAACATCGGCTTCCAGCTGATGATCACCATCGGCATCCTGGCGGCGGA 727

Qy        587 CGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGG 646
                ||   ||||| |   |  |   |   |          |||||||||| | |  || |
Db        728 GCTCATCAACTACGGCACCAACAAGATCAAGGCCGGGTACGGGTGGCGCGTGAGCCTGGC 787

Qy        647 CATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCC 706
              | |  |||| |||| | |   |  || | ||      |||  | | |||||  | ||
Db        788 GCTGGCGGCGGTGCCGGCGGCCATCATCACCCTGGGCTCCCTCTTCCTCCCGGACACCCC 847

Qy        707 GCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTC 766
                | |||| |   | |||  | ||  | ||| ||      ||||       | |
Db        848 CAACTCGCTGCTGGAGCGGGGCCACCCGGAGGAGGCACGCCGCATGCTCCGCCGCATCCG 907

Qy        767 CGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCC 826
              || |||||  ||    | || ||||  |  || |   |  ||||  ||  |   ||
Db        908 CGGCACGGACGACATCGGCGAGGAGTACGCGGACCTGGTGGCGGCCAGCGAGGAGGCCCG 967

Qy        827 TGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCG 886
              |||  ||  | ||
Db        968 CCAGGTGCGCCACCC---------------------------------------------- 982

Qy        887 GGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGG 946
              ||||  | |  ||  | |    ||  |  ||  ||   ||    |||          |
Db        983 ---GTGGCGGAACATCCTGCGCCGCCGGTACCGCGCGCAGCTCACCAT---------GGC 1030

Qy        947 GATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCC 1006
              |||  |||| ||||||||||||      ||| ||  ||  |  ||  | ||||||   ||
Db       1031 CGTCGCGATCCCCTTCTTCCAGCAGCTCAAGGGGATCAACGTCATCATGTTCTACGCGCC 1090

Qy       1007 TCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCC 1066
              |  ||||| | |  |  |||||  | |||  ||      | ||  ||  | ||       |
Db       1091 CGTGCTGTTCGACACGCTGGGATTCAAGAAAGAAGCCTTCCTCATGTCCTCCGTCATCAC 1150

Qy       1067 GTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGG 1126
              |   ||  ||  |||      |   || | |  || | |   ||  | ||||  ||||||
Db       1151 G---GGCCTCGTCAACGTCTTCGCCACCGTCGTGTCCATCGTCACCGTCGACCGCGTCGG 1207

Qy       1127 GCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGG 1186
               ||  ||   |||||| ||     ||   |   ||| ||| | | ||   ||  | | ||
Db       1208 CCGCCGCAAGCTGTTCCTCCAGGGCGGCGCGCAGATGATCGTGTGCCAGCTCATCGTGGG 1267

Qy       1187 CGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGC------CAAGATACCTTGGGCCAT 1240
              | |      || ||| |      |  |  |  || | |      ||          |   |
Db       1268 CACGCTCATCGCCGCCAAGTTCGGGACCAGCGGCACGGGGGACATCGCCAAGGGCTACGC 1327

Qy       1241 CGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCC 1300
              ||   |  | | | || || |  ||||||||||||  |||| ||          || ||
Db       1328 CGCGGTCGTCGTGGTCTTCATCTGCGCCTACGTCGCCGGCTTCGCCTGGTCGTGGGGGCC 1387

Qy       1301 CATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTC 1360
              | |    |||  |        || ||||||||||||||  || | ||| ||  |||
Db       1388 CCTGGGCTGGCTCGTGCCGTCCGAGATCTTCCCGCTGGAGATCCGCCCGGCGGGGCAGAG 1447

Qy       1361 GCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCT 1420
               ||  |||| ||| ||||    ||  || |  |  |||| |   | |||||||  |  |
Db       1448 CATCAACGTCTCCGTCAACATGTTCTTCACCTTCTGCATCGCGCAGGCCTTCCTCACCAT 1507

Qy       1421 GTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGC 1480
              |  |    | |||   |   ||||  ||||| | ||| | |||||   | ||   ||
Db       1508 GCTCTGCCACTTCAAGTT---CGGCCTCTTCTACTTCTTCGCCGGCTGGGTCGTCATCAT 1564

Qy       1481 CTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATG 1538
                 || ||  ||  |  || ||| || ||||||     |  | |  | |||||||||
Db       1565 GACCGTCTTTATCGCCTTCTTCCTGCCCGAGACCAAGAACGTGCCCATCGAGGAGATG 1622

RESULT 8
US-09-679-686B-11
; Sequence 11, Application US/09679686B
; Patent No. 6624343
```

; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen M.
; APPLICANT: Lightner, Jonathan E.
; APPLICANT: Rafalski, J. Antoni
; APPLICANT: Thorpe, Catherine J.
; TITLE OF INVENTION: HEXOSE CARRIER PROTEINS
; FILE REFERENCE: BB1160 US NA
; CURRENT APPLICATION NUMBER: US/09/679,686B
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: 60/081,131
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: PCT/US99/07561
; PRIOR FILING DATE: 1999-04-07
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 11
; LENGTH: 1776
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (66)
; OTHER INFORMATION: n = a, c, g or t
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (219)
; OTHER INFORMATION: n = a, c, g or t
US-09-679-686B-11

Query Match 8.3%; Score 168.4; DB 4; Length 1776;
Best Local Similarity 48.8%; Pred. No. 3.4e-24;
Matches 632; Conservative 1; Mismatches 597; Indels 66; Gaps 4;





RESULT 9
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match 7.3%; Score 147.2; DB 3; Length 4411529;

```
  Best Local Similarity   47.5%;  Pred. No. 4.5e-19;
  Matches  657;  Conservative    1;  Mismatches  634;  Indels   90;  Gaps    4;

Qy        156 GCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGG 215
              ||  ||  |  ||   ||  |||| |||   |  | |||||  ||| |||||  |
Db    3717236 GCGCTCCTCGTCGGGCTCACCGCCGCCAGCGTCGGCGTCCTCTACGGGTACGACCTTTCC 3717295

Qy        216 GTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTG 275
              |  ||    || |||| |||||   |||   |||| ||| |  |||       |  |   |
Db    3717296 GCCATCGCGGGTGCGTTGCTGTCTCTCAGCGAGGAATTCGAACTCACCACTCGAGAACAG 3717355

Qy        276 GAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGG 335
              ||| | ||  |    ||   |  |||  |  | ||  |||  |  |  |  |  ||
Db    3717356 GAGTTGCTGACCACCACGGCGGTGCTCGGCCAGATCGCCGGGGCGCTTGGCGGCGGCATC 3717415

Qy        336 ACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCG 395
                |  ||  |||||| ||   | |  |  | ||| ||  ||  ||  |   |  |  | |
Db    3717416 CTCGCCAACGCGATCGGACGCAAGAAATCGGTGGTGCTCATCGTCGCCGGCTACGCAGTG 3717475

Qy        396 GGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTG 455
                |:  | |||   | |  | | ||   |    | |||||  ||  | |  ||  | ||
Db    3717476 TTCGCCCTGCTCGGCGCGACCTCGGTGTCCGTACCGATGCTGGTGGTGGCGCGTCTGCTG 3717535

Qy        456 GCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCG 515
                ||  |    | |||    ||  |  ||  | | |||||||||    ||||||  |.||
Db    3717536 CTGGGTGTGACAATCGGCCTGTCGGTGGTGGTGGTGCCGGTGTATGTGGCCGAGTCGGCG 3717595

Qy        516 CCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATC 575
              ||||||  ||| ||||||  |   ||  | |          || ||   |       |||||||
Db    3717596 CCGGCGGCGGTGCGTGGGTCGTTGGTGACCGCGTATCAGCTGGCGACGCTTAGCGGCATC 3717655

Qy        576 CTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGC 635
                |  |||| ||| |      |||    ||  || | | ||           ||| ||||||
Db    3717656 GTCGTCGGTTACCTGGTCGGCTACCTGTTGGCCGGATCGCA---------CGGCTGGCGC 3717706

Qy        636 ATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATG 695
                |||  ||||   |  | ||| ||||||  ||  ||||||    ||   ||  ||  |||||
Db    3717707 GCGATGTTCGGGCTGGCCGCCGCGCCGGCCACGCTGCTGTTGCCGTTGTTGTGGCGCATG 3717766

Qy        696 CCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTG 755
              ||  ||   |  | ||  |||     |  | |||||  ||  |||| |||||        ||||
Db    3717767 CCCGATACCGCCCGCTGGTATCTGCTCAAGGGCCGGATCGCCGACGCGCGTAGCGCGCTG 3717826

Qy        756 GAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCC 815
                |  |                                              |||  ||  | |
Db    3717827 CGGCGG-----------------------------------------ATCCAGCCGGAG 3717844

Qy        816 GCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGA 875
              ||||  ||||  ||  |||||  | ||      | ||    | |||  | |  | ||  ||||
Db    3717845 GCCGACATCGATGCCGAGCTGGCCGATATGGCGGCCGCGGTCGACGAACGCGGCGGCGGT 3717904

Qy        876 AACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATC 935
              | ||   |  ||||  ||                        | | |   | |||   |
Db    3717905 ATCGGCGAAATGGTGCGG----------------------CGGCCGTATCTGCGGGCC 3717940

Qy        936 CTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTC 995
              |||||  ||  ||||   ||  |||| ||  ||      || ||  || |  || ||
Db    3717941 ACGCTGTTCGTCATCGCGCTCGGCTTCCTCGTCCAGATCACCGGGATCAACGCGATCATC 3718000

Qy        996 TTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGC 1055
              | ||||||  ||  |   | |||      || ||  ||  | |          | | | |
Db    3718001 TACTACAGTCCGCGACTTTTCGCCGCCATGGGCTTCGCGGGCTATTTCGCGATGCTTGCC 3718060

Qy       1056 ACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATC 1115
                |  | | |   ||| ||      |  | |    ||  ||  |  ||
Db    3718061 CTGCCCGCGATGGTGCAAGTCGCCGGCTTGGCGGCGGTGTGTGCCTCGCTGTTTCTGGTC 3718120

Qy       1116 GACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTC 1175
              ||  |  | || || || ||| |  || ||  | |||  |   ||| |||  | |
Db    3718121 GATCGGCTGGGCCGTCGCCCGATCCTGTTGTCCGGCATCGCGACGATGATCACCGCAGAT 3718180

Qy       1176 ATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTGG 1235
                ||  ||   | |||   || |                |  |||| || |       |
```

```
Db    3718181 GCCGTGCTGATCACCGTATTCGC--------------CAACGACTCCGATGGTGGCACG 3718225

Qy       1236 GCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTT 1295
                |  | |  |  | |||||  |   ||    | | ||  |  |||  | ||||  |
Db    3718226 GGGCTGGTGTTGGGGTTCGCCGGCGTGCTGCTGTTCATCATCGGGTTCAACTTCGGATTC 3718285

Qy       1296 GGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGC 1355
                ||| |  |  ||||||||  | |||||| |||||||  || | ||||  || ||||
Db    3718286 GGCTCGCTGGTCTGGGTGTACGCCGCGGAGAGCTTCCCGTCCCGGCTGCGGTCGATGGGA 3718345

Qy       1356 TGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTG 1415
                |  |  || || | |  |  |||  ||  ||  |  ||     ||  ||
Db    3718346 TCGAGCCCGATGCTCACCTCGACACTGACGGCCAACGCGATCGTTGCCGCCTTCTCGCTC 3718405

Qy       1416 TCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCG 1475
                |  ||      |   ||  |  ||  |||   |||    |||  |  |    ||||||
Db    3718406 ACCATGCTGCGTGTGCTCGGCGGCGCAGGCGTTTTCGCGGTCTTCGGCACGTTCGCCGTC 3718465

Qy       1476 CTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAG 1535
                ||||  |  |||  ||  |  |||   ||      |||||||||    |||||  |  || ||||||
Db    3718466 GTCGCGTTCGTGGTCGTGTACCGCTTTGCGCCGGAGACCAAGGGCCGCAAACTCGAGGAG 3718525

Qy       1536 AT 1537
              ||
Db    3718526 AT 3718527
```

```
RESULT 10
US-09-489-039A-4731
; Sequence 4731, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
;  APPLICANT:  Gary Breton et. al
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
;  TITLE OF INVENTION:  PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  2709.2004001
;  CURRENT APPLICATION NUMBER: US/09/489,039A
;  CURRENT FILING DATE:  2000-01-27
;  PRIOR APPLICATION NUMBER:  US 60/117,747
;  PRIOR FILING DATE:  1999-01-29
;  NUMBER OF SEQ ID NOS:  14342
; SEQ ID NO 4731
;   LENGTH: 1545
;   TYPE: DNA
;   ORGANISM: Klebsiella pneumoniae
US-09-489-039A-4731

  Query Match             7.2%;  Score 145.4;  DB 4;  Length 1545;
  Best Local Similarity   47.0%;  Pred. No. 1.1e-19;
  Matches  663;  Conservative    1;  Mismatches  672;  Indels   75;  Gaps    4;

Qy        127 CGCCGAAGAAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGA 186
                |||  ||  | |    |  | |  |||  |  ||          |||||  |||  |  ||
Db        158 CGCGGACGCAACGTGATACCCGGCGGATGAACTGGTTTGTTTCCATCGCCGCGGCGGTGG 217

Qy        187 CCTCCATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGA 246
                |  | | |||  | |||   |||||||| |||||   |||  |||| ||  |  || |
Db        218 CTGGCTTGCTCTTTGGCCTGGATATCGGCGTGATATCCGGAGCGTTGCCCTTTATAACCG 277

Qy        247 AGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACT 306
                | | ||||  |  |      |  |      ||||    |  |  |||   |||   |
Db        278 ACCATTTCACCTTATCCAGCCAGCTTCAGGAGTGGGTGGTTAGCAGCATGATGTTGGGGG 337

Qy        307 CGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCA 366
                ||  || |||  |  |      ||  ||  |||   | |  | |||||    |||| |
Db        338 CGGCGATAGGCGCGCTGTTTAACGGCTGGCTGTCTTTCCGCCTTGGCCGTAAATACAGCC 397

Qy        367 TCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACT 426
                |  ||   | |||||  | ||  | || ||  :  ||      |  || |||| || |
Db        398 TGATGGCGGGGGCCGTACTCTTTGTGGCCGGCTCTATCGGATCCGCTTTTGCCGCCAGCG 457

Qy        427 ACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGA 486
                  |  ||||  ||  |  | ||||  | |||   |||  |  |  ||  ||    ||
```

```
Db        458 TGGAGGTGCTGCTGGTGGCCCGCGTGGTGTTGGGCGTGGCAGTCGGGATTGCCTCTTATA 517

Qy        487 TCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGT 546

Db        518 CCGCGCCGCTGTACCTCTCCGAGATGGCCAGCGAGAACGTGCGCGGGAAAATGATCAGTA 577

Qy        547 TCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCT 606

Db        578 TGTATCAGCTGATGGTCACCCTTGGCATTGTGCTGGCGTTTCTTTCCGATACCGCCTTTA 637

Qy        607 CCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCG 666

Db        638 GCTAC---------AGCGGTAACTGGCGCGCCATGCTGGGCGTGCTGGCGCTGCCGGCGG 688

Qy        667 TGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGG 726

Db        689 TGATCCTGATCATTCTGGTCGTCTTTTTGCCGAACAGCCCGCGCTGGCTGGCGGAGAAGG 748

Qy        727 GACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCG 786

Db        749 GACGCCATATCGAAGCGGAAGAAGTGCTGCGGATGCTGCGCGATACCTCGGAAAAGGCGC 808

Qy        787 CGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACG 846

Db        809 GCGACGAGCTTAACGAGATC---------------------------------------- 828

Qy        847 TGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCC 906

Db        829 ----------------CGTGAGAGCCTCAAGCTGAAGCAGGGCGGTTGGGCGTTGTTTA 871

Qy        907 TGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCC 966

Db        872 AGATC---AATCGTAACGTGCGCCGGGCGGTGTTCCTGGGCATGCTGCTGCAGGCGATGC 928

Qy        967 AGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCG 1026

Db        929 AGCAATTCACCGGCATGAACATCATCATGTACTATGCGCCGCGTATCTTTAAAATGGCGG 988

Qy       1027 GATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGC 1086

Db        989 GCTTTACCACTACTGAACAGCAGATGATCGCCACCCTGGTGGTGGGCCTGACCTTTATGT 1048

Qy       1087 TTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGG 1146

Db       1049 TTGCCACCTTTATTGCGGTGTTCACAGTGGATAAAGCGGGCCGCAAGCCAGCGCTGAAAA 1108

Qy       1147 GCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCG 1206

Db       1109 TCGGCTTTAGCGTGATGGCGCTGGGCACTCTGGTGCTGGGCTACTGCCTGATGCAGTTCG 1168

Qy       1207 GCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCG 1266

Db       1169 ACAATGGCACCGCATCCA------GCGGCCTCTCCTGGCTCTCCGTCGGCATGACCATGA 1222

Qy       1267 CCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGA 1326

Db       1223 TGTGTATTGCCGGGTATGCGATGAGCGCGGCGCCGGTGGTGTGGATCCTCTGCTCCGAGA 1282

Qy       1327 TCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCA 1386

Db       1283 TCCAGCCGCTAAAATGCCGCGACTTCGGTATCACCTGCTCGACCACCACCAACTGGGTGT 1342

Qy       1387 CCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCA 1446

Db       1343 CGAACATGATCATCGGCGCCACTTTCCTGACGCTGCTTGACGCGATTGGCGCCGCCGGCA 1402

Qy       1447 GCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCC 1506

Db       1403 CCTTCTGGCTCTACACGGCGCTCAACGTGGCCTTTATCGGCATCACCTTCTGGCTGATCC 1462

Qy       1507 CGGAGACCCGCGGCCGGACGCTGGAGGAGAT 1537

Db       1463 CGGAAACCAAGAATGTCACCCTCGAGCACAT 1493
```

RESULT 11
US-09-679-686B-17
; Sequence 17, Application US/09679686B
; Patent No. 6624343
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen M.
; APPLICANT: Lightner, Jonathan E.
; APPLICANT: Rafalski, J. Antoni
; APPLICANT: Thorpe, Catherine J.
; TITLE OF INVENTION: HEXOSE CARRIER PROTEINS
; FILE REFERENCE: BB1160 US NA
; CURRENT APPLICATION NUMBER: US/09/679,686B
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: 60/081,131
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: PCT/US99/07561
; PRIOR FILING DATE: 1999-04-07
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 17
; LENGTH: 1752
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-679-686B-17

Query Match 7.0%; Score 141.2; DB 4; Length 1752;
Best Local Similarity 48.3%; Pred. No. 7.3e-19;
Matches 632; Conservative 1; Mismatches 609; Indels 66; Gaps 6;

```
Qy        237 TACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTG 296

Db        273 TCCGCCAGCCAGTACTGCAAGTACGACAACCAGCTGCTCCAGACCTTCACCTCCTCCCTC 332

Qy        297 AACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGG 356

Db        333 TACCTCGCGGCGCTCGTCTCCTCCTTCTTCGCCGCCACCGTCACACGTGTCGTGGGCCGT 392

Qy        357 CGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGSGTTCCTCATGGGGTTC 416

Db        393 AAGTGGTCCATGTTCACCGGAGGGCTCACCTTCCTCATCGGCGCTGCGCTTAACGGGGCG 452

Qy        417 GCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTAC 476

Db        453 GCGGAGAACATCGCCATGCTCATCGTCGGACGCATCCTCCTCGGTGTCGGCGTTGGCTTC 512

Qy        477 GCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTC 536

Db        513 GCCAATCAGTCTGTGCCGGTGTACCTGTCGGAGATGGCGCCTGCGCGTCTCCGGGGCATG 572

Qy        537 CTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAAC 596

Db        573 CTCAACATCGGGTTCCAGCTCATGATCACCATCGGCATCCTGGCGGCGGCGCTCATCAAT 632

Qy        597 TATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCG 656

Db        633 TACGACACCAACAAGATCAAGGCCGGGTACGGGTGGCGCATCAGCCTGGCCATCGCGGCC 692

Qy        657 GCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTG 716

Db        693 GTCCCGGCGGGCATCATCACCCTGGGGTCCTTTTTCCTCCCCGACACCCCCAACTCCCTC 752

Qy        717 GTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCG 776

Db        753 ATCGAGCGTGGCCACCCGGAGGCGGCGCGCCGCATGCTCAACCGCATCCGCGGCAGCGAC 812

Qy        777 GAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTC 836

Db        813 GTGGA----------------------------------------CATCAGCGAGGAGTAC 833

Qy        837 GACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAG 896
```

```
Db        834 GCGGACCTGGTGGTG---------------GCGAGCGAGGAGTCCAAGCTGGTGCAGCAC 878

Qy        897 GAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATC 956

Db        879 CCGTGGCGCAACATCCTGCAGCGCAAGTACCGGCCCCAGCTGACCATGGCGATCATGATC 938

Qy        957 CACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTC 1016

Db        939 CCCTTCTTCCAGCAGCTGACGGGCATCAACGTCATCATGTTCTACGCGCCGGTGCTGTTC 998

Qy       1017 AAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTC 1076

Db        999 GAGACGCTGGGGTTCAAGGGCGACGCGTCGCTCATGTCGGCCGTCATCACGGGCCTGGTC 1058

Qy       1077 ACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCG 1136

Db       1059 AACGTGTTCGCGACGCTCGTGTCCGTG---TTCACCGTCGACCGGCTGGGTCGCCGGAAG 1115

Qy       1137 CTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTC 1196

Db       1116 CTGTTCCTGCAGGGCGGCACGCAGATGCTGCTGAGCCAGCTGGTGGTGGGCACCCTGATC 1175

Qy       1197 ACCGTC---GTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATC 1253

Db       1176 GCGGTCAAGTTCGGGACGAGCGGCGTGGGGGAGATGCCCAAGGGGTACGCGGCGGCGGTG 1235

Qy       1254 GC---CTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGG 1310

Db       1236 GTGCTCTTCATCTGCCTCTATGTGGCCGGGTTCGCGTGGTCGTGGGGGCCCCTGGGGTGG 1295

Qy       1311 GTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTC 1370

Db       1296 CTGGTGCCCAGCGAGATCTTCCCGCTGGAGATCAGGCCGGCGGGGCAGAGCATCAACGTG 1355

Qy       1371 GCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCC 1430

Db       1356 TCGGTGAACATGCTCTTCACCTTCGTCATCGCGCAGGCGTTCCTCACCATGCTCTGCCAC 1415

Qy       1431 ATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTC 1490

Db       1416 ATGAAGTT---CGGCCTCTTCTACTTCTTCGCCGGCTGGGTGGTGATCATGACCGTCTTC 1472

Qy       1491 TTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATG 1538

Db       1473 ATCGCGCTCTTCCTGCCGGAGACCAAGAACGTGCCCATCGAGGAGATG 1520


RESULT 12
US-09-679-686B-9
; Sequence 9, Application US/09679686B
; Patent No. 6624343
; GENERAL INFORMATION:
;  APPLICANT: Allen, Stephen M.
;  APPLICANT:  Lightner, Jonathan E.
;  APPLICANT:  Rafalski, J. Antoni
;  APPLICANT:  Thorpe, Catherine J.
;  TITLE OF INVENTION: HEXOSE CARRIER PROTEINS
;  FILE REFERENCE: BB1160 US NA
;  CURRENT APPLICATION NUMBER: US/09/679,686B
;  CURRENT FILING DATE:  2003-01-16
;  PRIOR APPLICATION NUMBER: 60/081,131
;  PRIOR FILING DATE: 1998-04-09
;  PRIOR APPLICATION NUMBER: PCT/US99/07561
;  PRIOR FILING DATE: 1999-04-07
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 9
;   LENGTH: 1167
;   TYPE: DNA
;   ORGANISM: Oryza sativa
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1020)
;   OTHER INFORMATION: n = a, c, g or t
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1031)
;   OTHER INFORMATION: n = a, c, g or t
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1059)
;   OTHER INFORMATION: n = a, c, g or t
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1086)
;   OTHER INFORMATION: n = a, c, g or t
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1126)
;   OTHER INFORMATION: n = a, c, g or t
US-09-679-686B-9

  Query Match             6.7%;  Score 136;  DB 4;  Length 1167;
  Best Local Similarity   51.5%;  Pred. No. 6.8e-18;
  Matches  310;  Conservative    1;  Mismatches  291;  Indels    0;  Gaps    0;

Qy        218 GATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGA 277

Db        323 GAAGAAGCAGATGGCGGACAAGAACAACCAGTACTGCAAGTACGACAACCAGCTGCTGCA 382

Qy        278 GGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGAC 337

Db        383 GACCTTCACCTCGTCGCTCTACCTCGCCGCCCTCGTCTCCTCCTTCTTCGCCGCCACCGT 442

Qy        338 GTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGG 397

Db        443 CACCCGCGTCCTCGGCCGCAAGTGGTCCATGTTCGCCGGCGGCCTCACCTTCCTCATCGG 502

Qy        398 GGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGC 457

Db        503 CGCCGCCCTCAACGGCGCCGCCGAGAACGTCGCCATGCTCATCGTCGGTCGTATCCTCCT 562

Qy        458 CGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCC 517

Db        563 CGGTGTCGGCGTCGGCTTCGCCACTCAGTCGGTGCCGGTGTACTTGTCGGAGATGGCGCC 622

Qy        518 GGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCT 577

Db        623 GGCTCGGCTGCGGGGGATGCTGAACATCGGGTTCCAGCTGATGATCACCATCGGCATCCT 682

Qy        578 GCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCAT 637

Db        683 GGCGGCGGAGCTGATAAACTACGGGACGGCGAAGATCAAGGCCGGGTGGGGATGGCGGGT 742

Qy        638 CATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCC 697

Db        743 GAGCCTGGCGCTGGCCGCCGTCCCCGCCGCCATCATCACCCTCGGCTCCCTCTTCCTCCC 802

Qy        698 GGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGA 757

Db        803 GGACACCCCCAACTCGCTCATCGACAGGGGCCACCCGGAGGCGGCGGAGCGCATGCTCCG 862

Qy        758 GAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGC 817

Db        863 GCGCATCCGCGGCTCCGACGTGGACGTGTCGGAGGAGTACGCGGACCTGGTGGCGGCGAG 922

Qy        818 CG 819

Db        923 CG 924


RESULT 13
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
```

```
;  APPLICANT: FLEISCHMAN, Robert D.
;  APPLICANT:  WHITE, Owen R.
;  APPLICANT:  FRASER, Claire M.
;  APPLICANT:  VENTER, John C.
;  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
;  TITLE OF INVENTION:  TUBERCULOSIS
;  FILE REFERENCE: 24366-20007.00
;  CURRENT APPLICATION NUMBER: US/09/103,840A
;  CURRENT FILING DATE:  1998-06-24
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 4403765
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
;   FEATURE:
;   OTHER INFORMATION: CDC 1551
;   OTHER INFORMATION: "n" bases at various positions throughout the sequence
;   OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

  Query Match             6.3%;  Score 128;  DB 3;  Length 4403765;
  Best Local Similarity   47.7%;  Pred. No. 2.6e-15;
  Matches  659;  Conservative    1;  Mismatches  626;  Indels   96;  Gaps    6;

Qy        156 GCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGG 215
              ||  ||  |  ||   ||  |||| |||    |  | |||||  ||| |||||  |
Db    3714802 GCGCTCCTCGTCGGGCTCACCGCCGCCAGCGTCGGCGTCCTCTACGGGTACGACCTTTCC 3714861

Qy        216 GTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTG 275
              |  ||    || |||| |||||   |||   |||| ||| |  |||       |  |   |
Db    3714862 GCCATCGCGGGTGCGTTGCTGTCTCTCAGCGAGGAATTCGAACTCACCACTCGAGAACAG 3714921

Qy        276 GAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGG 335
              ||| |  || |     ||    |  |||  |  | ||  |||  |  | |  |  ||
Db    3714922 GAGTTGCTGACCACCACGGCGGTGCTCGGCCAGATCGCCGGGGCGCTTGGCGGCGGCATC 3714981

Qy        336 ACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCG 395
              |  ||  |||||| ||  | | |  | ||| ||  || || |   | |   | |
Db    3714982 CTCGCCAACGCGATCGGACGCAAGAAATCGGTGGTGCTCATCGTCGCCGGCTACGCAGTG 3715041

Qy        396 GGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTG 455
              |:  | |||   | |  | | ||   |    | |||||  || | |  ||  |  ||
Db    3715042 TTCGCCCTGCTCGGCGCGACCTCGGTGTCCGTACCGATGCTGGTGGTGGCGCGTCTGCTG 3715101

Qy        456 GCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCG 515
              ||  |     | |||   || |  || | | |||||||||    ||||||  | ||
Db    3715102 CTGGGTGTGACAATCGGCCTGTCGGTGGTGGTGGTGCCGGTGTATGTGGCCGAGTCGGCG 3715161

Qy        516 CCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATC 575
              |||||| ||| |||||| |   ||  | |       || ||   |       |||||||
Db    3715162 CCGGCGGCGGTGCGTGGGTCGTTGGTGACCGCGTATCAGCTGGCGACGCTTAGCGGCATC 3715221

Qy        576 CTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGC 635
              |  |||| ||| |       |||    ||  || | | ||          ||| ||||||
Db    3715222 GTCGTCGGTTACCTGGTCGGCTACCTGTTGGCCGGATCGCA---------CGGCTGGCGC 3715272

Qy        636 ATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATG 695
              |||  ||||  | | ||| ||||||  ||  ||||||   ||   || ||    |||||
Db    3715273 GCGATGTTCGGGCTGGCCGCCGCGCCGGCCACGCTGCTGTTGCCGTTGTTGTGGCGCATG 3715332

Qy        696 CCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTG 755
              |||    |||| || |||     |  | ||||| ||  |||| |||||
Db    3715333 CCGATACCGCC-CGCTGGTATCTGCTCAAGGGCCGGATCGCCGACGCG------------ 3715379

Qy        756 GAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCC 815
                  ||   | |||| |||  | || |  ||
Db    3715380 -----------CGTAGCTGCGGCGGATCCAGCCGGACG---------------------- 3715406

Qy        816 GCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGA 875
              || ||||  ||  ||||| | ||     | ||   | |||  | |  | ||  ||||
Db    3715407 --CGACATCGATGCCGAGCTGGCCGATATGGCGGCCGCGGTCGACGAACGCGGCGGCGGT 3715464

Qy        876 AACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATC 935
              | ||   |  |||| ||  |  |   | | | || |
Db    3715465 ATCGGCGAAATGGTGCGGCGGCCGTATCTGCGGGCCAC---------------------- 3715502

Qy        936 CTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTC 995
              |||||  ||  ||||   ||  |||| ||   ||       ||  ||  ||  |  || ||
Db    3715503 --GCTGTTCGTCATCGCGCTCGGCTTCCTCGTCCAGATCACCGGGATCAACGCGATCATC 3715560

Qy        996 TTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGC 1055
              | ||||||  ||  |   | |||     |   || ||  ||  | |       |  | | |
Db    3715561 TACTACAGTCCGCGACTTTTCGCCGCCATGGGCTTCGCGGGCTATTTCGCGATGCTTGCC 3715620

Qy       1056 ACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATC 1115
                  |   | | |   ||| ||     |       | |       ||   ||  | ||
Db    3715621 CTGCCCGCGATGGTGCAAGTCGCCGGCTTGGCGGCGGTGTGTGCCTCGCTGTTTCTGGTC 3715680

Qy       1116 GACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTC 1175
              ||  |  | || ||  ||  ||| |  || ||  | |||  |    ||| |||  | |
Db    3715681 GATCGGCTGGGCCGTCGCCCGATCCTGTTGTCCGGCATCGCGACGATGATCACCGCAGAT 3715740

Qy       1176 ATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTGG 1235
              ||  ||   | |||   || |              |  |||| || |           |
Db    3715741 GCCGTGCTGATCACCGTATTCGC--------------CAACGACTCCGATGGTGGCACG 3715785

Qy       1236 GCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTT 1295
              |  | |   |  |  |||||  |  ||   | | ||  |   |||  | ||||  |
Db    3715786 GGGCTGGTGTTGGGGTTCGCCGGCGTGCTGCTGTTCATCATCGGGTTCAACTTCGGATTC 3715845

Qy       1296 GGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGC 1355
              |||  |   |||||||||  | ||||||  |||||||   ||  | ||||  || ||||
Db    3715846 GGCTCGCTGGTCTGGGTGTACGCCGCGGAGAGCTTCCCGTCCCGGCTGCGGTCGATGGGA 3715905

Qy       1356 TGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTG 1415
              |    ||     ||  || | |  |.   ||| ||   || |  ||    ||  ||
Db    3715906 TCGAGCCTGATGCTCACCTCGACACTGACGGCCAACGCGATCGTTGCCGCCTTCTCGCTC 3715965

Qy       1416 TCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCG 1475
              |  ||      |   ||  |  ||  |||    |||    ||| |  |    ||||||
Db    3715966 ACCATGCTGCGTGTGCTCGGCGGCGCAGGCGTTTTCGCGGTCTTCGGCACGTTCGCCGTC 3716025

Qy       1476 CTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAG 1535
              ||||  |  |||  ||  |  |||  ||      |||||||||    |||||  |  ||  ||||||
Db    3716026 GTCGCGTTCGTGGTCGTGTACCGCTTTGCGCCGGAGACCAAGGGCCGCAAACTCGAGGAG 3716085

Qy       1536 AT 1537
              ||
Db    3716086 AT 3716087


RESULT 14
US-09-489-039A-4762
; Sequence 4762, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
;  APPLICANT:  Gary Breton et. al
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
;  TITLE OF INVENTION:  PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  2709.2004001
;  CURRENT APPLICATION NUMBER: US/09/489,039A
;  CURRENT FILING DATE:  2000-01-27
;  PRIOR APPLICATION NUMBER:  US 60/117,747
;  PRIOR FILING DATE:  1999-01-29
;  NUMBER OF SEQ ID NOS:  14342
; SEQ ID NO 4762
;   LENGTH: 1431
;   TYPE: DNA
;   ORGANISM: Klebsiella pneumoniae
US-09-489-039A-4762

  Query Match             5.9%;  Score 118;  DB 4;  Length 1431;
  Best Local Similarity   50.5%;  Pred. No. 2.4e-14;
  Matches  317;  Conservative    1;  Mismatches  301;  Indels    9;  Gaps    1;
```

```
Qy        157 CCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGGG 216
Db         80 CGTTCTTCGTCTGTTTCCTCGCCGCGCTGGCTGGCCTGCTGTTCGGCCTTGATATCGGTG 139

Qy        217 TGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGG 276
Db        140 TTATTGCGGGTGCCTTACCCTTTATTGCCAATGAGTTCCAGATTTCCGCCCACACCCAGG 199

Qy        277 AGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGA 336
Db        200 AGTGGGTGGTCAGCTCCATGATGTTCGGGGCTGCCGTCGGCGCGGTCGGCAGCGGCTGGC 259

Qy        337 CGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGG 396
Db        260 TCTCTTTCAAACTGGGCCGGAAAAAGAGCCTGATGATCGGCGCCATCCTCTTCGTCGCCG 319

Qy        397 GGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGG 456
Db        320 GTTCGCTGTTCTCTGCCGCCGCGCCAAACGTCGAGATCCTGCTGGTTTCCCGTGTGCTGC 379

Qy        457 CCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGC 516
Db        380 TCGGCCTGGCGGTGGGCGTCGCCTCATATACGGCTCCGCTGTATCTGTCGGAAATCGCCC 439

Qy        517 CGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCC 576
Db        440 CAGAAAAAATTCGCGGCAGTATGATTTCCATGTACCAGCTGATGATCACCATCGGGATCC 499

Qy        577 TGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCA 636
Db        500 TTGGCGCCTATCTCTCTGACACCGCTTTCAGCTAC---------AGCGGCGCATGGCGCT 550

Qy        637 TCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGC 696
Db        551 GGATGCTCGGGGTTATCATCATTCCGGCGGTTTTGCTGCTGATCGGCGTTATCTTCCTGC 610

Qy        697 CGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGG 756
Db        611 CGGACAGCCCGCGCTGGTTCGCCGCCAAACGTCGCTTTGTCGATGCGGAACGCGTGCTGC 670

Qy        757 AGAAGACCTCCGACACGGCGGAGGAGGC 784
Db        671 TGCGCCTGCGCGATACCAGCGCCGAAGC 698
```

RESULT 15
US-09-252-991A-3992
; Sequence 3992, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 3992
; LENGTH: 1344
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (870)
; OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.
US-09-252-991A-3992

Query Match 5.1%; Score 102; DB 4; Length 1344;
Best Local Similarity 48.0%; Pred. No. 3.3e-11;

Matches 288; Conservative 1; Mismatches 311; Indels 0; Gaps 0;

```
Qy        169 CCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGGG 228
Db         77 CCCTGCGCTCCGGCGGCGGCTCCTTCGTCGACGGCTACGTGCTGAGCATCATCGGCGTGG 136

Qy        229 CGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGG 288
Db        137 CGATGCTGCAGATGACCGCGGCATTGCAACTGAGCAGTTTCTGGCAGGGCATGATCGCCG 196

Qy        289 GCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGA 348
Db        197 CCTCGGCGCTGGTCGGCATCTTCTTCGGCGGTTTCCTCGGCGGCTGGCTGACCGACCGCC 256

Qy        349 TCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGSGTTCCTCA 408
Db        257 TGGGCCGCCGCAGGCTGTTCTTCGTCGGCCCCAGCCTGTTCGTCCTCGCTTCCCTGGCGC 316

Qy        409 TGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCG 468
Db        317 AGTACGGGGTCGACTCGGCGGCGGCGCTGTTCGTCCTGCGCTTCCTGATCGGCGTGGCGG 376

Qy        469 TGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGC 528
Db        377 TGGGCATCGAGTACCCGGTCGCCACCGCCTTGCTGGTGGAGTTCCTGCCGAAGAAGTATC 436

Qy        529 GTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACG 588
Db        437 GCGGACCGCGCCTGGCCACCCTGACCATCCTCTGGTTCGCCGGCGCCGCCTTCGCCTACC 496

Qy        589 TCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCA 648
Db        497 TGGTCGGCGACCTGATCCTGCGCAGCGGCGGCGAGGACGCCTGGCGCCTGGTACTGGCCA 556

Qy        649 TCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGC 708
Db        557 GCACCGCGGCGATCGGTGCCGTGCTGTTCCTGCTACGCATCGGCACCCCCGAGTCGCCGC 616

Qy        709 GGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCG 768
Db        617 GCTGGCTGCTGAGCAAGGGCCGCGACGCGGACGCAGAGCGGGTCATCCGCCAGGTCTACG 676
```

Search completed: July 2, 2004, 08:30:48
Job time : 199 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: July 2, 2004, 06:17:31 ; Search time 4038 Seconds
(without alignments)
2409.714 Million cell updates/sec

Title: US-10-051-902A-21
Perfect score: 2017
Sequence: 1 cttacatgtaagctcgtgcc.....aaaaaaaaaaaaaaaaaaaa 2017

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3163042 seqs, 2412103800 residues

Total number of hits satisfying chosen parameters: 6326084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2016.6	100.0	2017	14	US-10-051-902-21	Sequence 21, Appl
2	2016.6	100.0	2017	14	US-10-051-909-21	Sequence 21, Appl
3	1382.8	68.6	1897	17	US-10-437-963-87990	Sequence 87990, A
4	1094.6	54.3	2089	14	US-10-051-902-25	Sequence 25, Appl
5	1094.6	54.3	2089	14	US-10-051-909-25	Sequence 25, Appl
6	1039.8	51.6	1093	13	US-10-425-114-30198	Sequence 30198, A
7	1005	49.8	1914	14	US-10-051-902-19	Sequence 19, Appl
8	1005	49.8	1914	14	US-10-051-909-19	Sequence 19, Appl
9	981.6	48.7	1872	14	US-10-051-902-27	Sequence 27, Appl
10	981.6	48.7	1872	14	US-10-051-909-27	Sequence 27, Appl
11	779.6	38.7	1704	13	US-10-425-114-28112	Sequence 28112, A
12	759.4	37.6	1809	17	US-10-437-963-78515	Sequence 78515, A
13	739.6	36.7	2203	13	US-10-425-114-30135	Sequence 30135, A
14	738	36.6	2231	13	US-10-425-114-26589	Sequence 26589, A
15	734.8	36.4	1777	13	US-10-425-114-28469	Sequence 28469, A
16	723.6	35.9	1404	13	US-10-425-114-3003	Sequence 3003, Ap
17	697.4	34.6	1220	13	US-10-425-114-28195	Sequence 28195, A
18	688.4	34.1	2362	17	US-10-437-963-73071	Sequence 73071, A
19	677.4	33.6	1073	13	US-10-425-114-23054	Sequence 23054, A
20	657.4	32.6	1722	17	US-10-437-963-31143	Sequence 31143, A
21	634.6	31.5	1275	13	US-10-425-114-35317	Sequence 35317, A
22	585	29.0	1685	17	US-10-437-963-48312	Sequence 48312, A
23	576.2	28.6	1720	13	US-10-425-114-12666	Sequence 12666, A
24	575.2	28.5	1853	14	US-10-051-902-23	Sequence 23, Appl
25	575.2	28.5	1853	14	US-10-051-909-23	Sequence 23, Appl
26	575	28.5	1485	17	US-10-437-963-92696	Sequence 92696, A
27	568.8	28.2	1437	17	US-10-437-963-29421	Sequence 29421, A
28	567	28.1	1426	13	US-10-425-114-16979	Sequence 16979, A
29	546.4	27.1	1608	17	US-10-437-963-62573	Sequence 62573, A
30	530.6	26.3	1214	13	US-10-425-114-5535	Sequence 5535, Ap
31	512.6	25.4	1843	17	US-10-437-963-22190	Sequence 22190, A
32	508.4	25.2	1450	13	US-10-424-599-57033	Sequence 57033, A
33	486	24.1	1429	13	US-10-425-114-32794	Sequence 32794, A
C 34	437.6	21.7	4014	13	US-10-424-599-56321	Sequence 56321, A
35	433.8	21.5	1482	9	US-09-938-842A-795	Sequence 795, App
36	433.8	21.5	1482	11	US-09-938-842A-795	Sequence 795, App
37	410.4	20.3	1404	13	US-10-425-114-30458	Sequence 30458, A
38	408	20.2	1669	13	US-10-424-599-26761	Sequence 26761, A
39	407.2	20.2	1507	13	US-10-425-114-9962	Sequence 9962, Ap
40	367.6	18.2	1326	17	US-10-437-963-98338	Sequence 98338, A
41	367.2	18.2	860	17	US-10-437-963-55632	Sequence 55632, A
42	353.4	17.5	1473	17	US-10-437-963-62571	Sequence 62571, A
43	343.6	17.0	615	14	US-10-051-902-17	Sequence 17, Appl
44	343.6	17.0	615	14	US-10-051-909-17	Sequence 17, Appl
45	341.6	16.9	2051	16	US-10-310-154-356	Sequence 356, App

ALIGNMENTS

RESULT 1
US-10-051-902-21
; Sequence 21, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 21
; LENGTH: 2017
; TYPE: DNA
; ORGANISM: Oryza sativa
US-10-051-902-21

Query Match 100.0%; Score 2016.6; DB 14; Length 2017;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2017; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CTTACATGTAAGCTCGTGCCGGCACGAGCTTACACTCGACCGCCACTACTGTACACGGCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTTACATGTAAGCTCGTGCCGGCACGAGCTTACACTCGACCGCCACTACTGTACACGGCC 60

Qy         61 CAGAGCGAGCCTCCTCCTCCTCTGCACCACCGGAGATGGCTTCCGCCGCGCTGCCGGAGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGAGCGAGCCTCCTCCTCCTCTGCACCACCGGAGATGGCTTCCGCCGCGCTGCCGGAGG 120

Qy        121 CCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        121 CCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCT 180

Qy        181 CCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACA 240

Qy        241 TCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACC 300

Qy        301 TCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGT 360

Qy        361 ACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCG 420

Qy        421 TCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGC 480

Qy        481 TCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGA 540

Qy        541 CGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATG 600

Qy        601 CTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGC 660

Qy        661 CGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCA 720

Qy        721 TGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGG 780

Qy        781 AGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACG 840

Qy        841 GCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGC 900

Qy        901 TCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACT 960

Qy        961 TCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGA 1020

Qy       1021 GCCCCGGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GCCCCGGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCA 1080

Qy       1081 AGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGT 1140

Qy       1141 TGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCG 1200

Qy       1201 TCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCA 1260

Qy       1261 CCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCT 1320

Qy       1321 CGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACC 1380

Qy       1381 GCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 GCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCG 1440

Qy       1441 GCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCT 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCT 1500

Qy       1501 ACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 ACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGG 1560

Qy       1561 CCGCCGCCTCGGAATCAGACGAGCCAGCCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCA 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CCGCCGCCTCGGAATCAGACGAGCCAGCCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCA 1620

Qy       1621 CTAACTGATCAAACTAACCGCAAAATCACCAAATCCTAAGGGTTTTCTTGCAAAAACGTG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 CTAACTGATCAAACTAACCGCAAAATCACCAAATCCTAAGGGTTTTCTTGCAAAAACGTG 1680

Qy       1681 TGCTGTACTGGCTAGCTAGCAAGTAGTAGCAGCAACGTGGGAAGATTCGCTGATCCGGCG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 TGCTGTACTGGCTAGCTAGCAAGTAGTAGCAGCAACGTGGGAAGATTCGCTGATCCGGCG 1740

Qy       1741 TTGCTGGAGAGCGACGGCCGGCGACGACAAAGCTGAGCTCCAGCTCGAGACTTCTTAAAA 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TTGCTGGAGAGCGACGGCCGGCGACGACAAAGCTGAGCTCCAGCTCGAGACTTCTTAAAA 1800

Qy       1801 TCATCTTCAAGTACATGGATTTTATTTTGCTCTTTGCTTTGTCCGTAAAAGTTGTACTAT 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 TCATCTTCAAGTACATGGATTTTATTTTGCTCTTTGCTTTGTCCGTAAAAGTTGTACTAT 1860

Qy       1861 GCGATGAAGAATACCAGTATGTAGCAAGGCTGAGGTTGTGTGTAGCTACTAGAAGTGTCA 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 GCGATGAAGAATACCAGTATGTAGCAAGGCTGAGGTTGTGTGTAGCTACTAGAAGTGTCA 1920

Qy       1921 GTCACGTTGTTCTTGTAAGAAATGTTTAACTGTTAATTAAGCAGTATTGTTGCAGTAATC 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GTCACGTTGTTCTTGTAAGAAATGTTTAACTGTTAATTAAGCAGTATTGTTGCAGTAATC 1980

Qy       1981 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2017
              |||||||||||||||||||||||||||||||||||||||||||
Db       1981 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2017
```

RESULT 2
US-10-051-909-21
; Sequence 21, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB1163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 21

; LENGTH: 2017
; TYPE: DNA
; ORGANISM: Oryza sativa
US-10-051-909-21

Query Match 100.0%; Score 2016.6; DB 14; Length 2017;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2017; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CTTACATGTAAGCTCGTGCCGGCACGAGCTTACACTCGACCGCCACTACTGTACACGGCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTTACATGTAAGCTCGTGCCGGCACGAGCTTACACTCGACCGCCACTACTGTACACGGCC 60

Qy         61 CAGAGCGAGCCTCCTCCTCCTCTGCACCACCGGAGATGGCTTCCGCCGCGCTGCCGGAGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGAGCGAGCCTCCTCCTCCTCTGCACCACCGGAGATGGCTTCCGCCGCGCTGCCGGAGG 120

Qy        121 CCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCT 180

Qy        181 CCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACA 240

Qy        241 TCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACC 300

Qy        301 TCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGT 360

Qy        361 ACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGGSGTTCCTCATGGGGTTCGCCG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGGSGTTCCTCATGGGGTTCGCCG 420

Qy        421 TCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGC 480

Qy        481 TCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGA 540

Qy        541 CGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATG 600

Qy        601 CTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGC 660

Qy        661 CGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCA 720

Qy        721 TGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGG 780

Qy        781 AGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACG 840

Qy        841 GCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGC 900

Qy        901 TCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACT 960

Qy        961 TCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGA 1020

Qy       1021 GCCCCGGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GCCCCGGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCA 1080

Qy       1081 AGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGT 1140

Qy       1141 TGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCG 1200

Qy       1201 TCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCA 1260

Qy       1261 CCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCT 1320

Qy       1321 CGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACC 1380

Qy       1381 GCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 GCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCG 1440

Qy       1441 GCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCT 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCT 1500

Qy       1501 ACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 ACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGG 1560

Qy       1561 CCGCCGCCTCGGAATCAGACGAGCCAGCCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCA 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CCGCCGCCTCGGAATCAGACGAGCCAGCCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCA 1620

Qy       1621 CTAACTGATCAAACTAACCGCAAAATCACCAAATCCTAAGGGTTTTCTTGCAAAAACGTG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 CTAACTGATCAAACTAACCGCAAAATCACCAAATCCTAAGGGTTTTCTTGCAAAAACGTG 1680

Qy       1681 TGCTGTACTGGCTAGCTAGCAAGTAGTAGCAGCAACGTGGGAAGATTCGCTGATCCGGCG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 TGCTGTACTGGCTAGCTAGCAAGTAGTAGCAGCAACGTGGGAAGATTCGCTGATCCGGCG 1740

Qy       1741 TTGCTGGAGAGCGACGGCCGGCGACGACAAAGCTGAGCTCCAGCTCGAGACTTCTTAAAA 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TTGCTGGAGAGCGACGGCCGGCGACGACAAAGCTGAGCTCCAGCTCGAGACTTCTTAAAA 1800

Qy       1801 TCATCTTCAAGTACATGGATTTTATTTTGCTCTTTGCTTTGTCCGTAAAAGTTGTACTAT 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 TCATCTTCAAGTACATGGATTTTATTTTGCTCTTTGCTTTGTCCGTAAAAGTTGTACTAT 1860

Qy       1861 GCGATGAAGAATACCAGTATGTAGCAAGGCTGAGGTTGTGTGTAGCTACTAGAAGTGTCA 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 GCGATGAAGAATACCAGTATGTAGCAAGGCTGAGGTTGTGTGTAGCTACTAGAAGTGTCA 1920

Qy       1921 GTCACGTTGTTCTTGTAAGAAATGTTTAACTGTTAATTAAGCAGTATTGTTGCAGTAATC 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GTCACGTTGTTCTTGTAAGAAATGTTTAACTGTTAATTAAGCAGTATTGTTGCAGTAATC 1980

Qy       1981 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2017
              ||||||||||||||||||||||||||||||||||||||||
Db       1981 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2017
```

RESULT 3
US-10-437-963-87990
; Sequence 87990, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 87990
; LENGTH: 1897
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_86885C.1
US-10-437-963-87990

Query Match 68.6%; Score 1382.8; DB 17; Length 1897;
Best Local Similarity 84.5%; Pred. No. 0;
Matches 1587; Conservative 1; Mismatches 263; Indels 27; Gaps 2;

```
Qy        135 AAGAAGGGCAACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATC 194
              ||||||   |||| | | |  || ||| ||||||||||||||||||||||||| ||||||||
Db         16 AAGAAGAAGAACGCCAGCTATGCATTCACCTGCGCCATCCTCGCCTCCATGGCCTCCATC 75

Qy        195 CTCCTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTC 254
              | ||||||| || |||||||||||||||| || |||||||||||||||| ||| |
Db         76 ATTCTCGGCTATGACATCGGGGTGATGAGCGGCGCATCGCTGTACATCAAGAAAGACCTG 135

Qy        255 AACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATC 314
              || |||| ||||| | |||||||| | || |||||||| |||||| |||||||||| ||
Db        136 AAGATCACTGACGTGCAGGTGGAGATCCTGATGGGCATCCTGAACATCTACTCGCTCGTC 195

Qy        315 GGCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTC 374
              |||||||||| |||||| ||||| |||||||||||||||||||||| |||| |||| |||
Db        196 GGCTCCTTCGCCGCGGGGAGGACGGCGGACTGGATCGGCCGGCGGTTCACCGTCGTCTTC 255

Qy        375 GCCGCCGTCATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATG 434
              || |||| | | |||||||| | |: | |||||||||||| ||| |||||||||
Db        256 GCTGCCGCCTTCTTCTTCGCCAGAGCTCTACTCATGGGGTTCTCCGGTGACTACGCCACT 315

Qy        435 CTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCG 494
              |||||| |||||||||||||||| ||| |||||||| |||||| ||||||||||||
Db        316 CTCATGGTCGGCCGCTTCGTGGCGGGCGTCGGCGTGGGGTACGCGATCATGATCGCGCCC 375

Qy        495 GTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAG 554
              |||||||| |||||| | |||||||||||| |||||||||||| ||||||||||||||||
Db        376 GTGTACACGGCCGAGATCTCGCCGGCGTCGTCGCGTGGCTTCCTCACGTCGTTCCCGGAG 435

Qy        555 GTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTG 614
              || | ||||||| |||| ||||| ||||| |||||||| |||||||||||| | ||| ||
Db        436 GTCTCCATCAACCTCGGTATCCTCCTCGGCTACGTCTCCAACTATGCTTTCGCGCGCCTG 495

Qy        615 CCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTC 674
              |||||   |||||||||||| |||||||||| |||||||||||| |||||||||||||||
Db        496 CCGCTCTCCCTCGGGTGGCGCGTCATGCTCGGCGTCGGCGCGGCGCCATCCGTGCTGCTC 555

Qy        675 GCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTC 734
              |||||||||||||||||||||||||||||||||||||||| |||||||| |||||||||
Db        556 GCGCTCATGGTGCTCGGCATGCCGGAGTCGCCACGGTGGCTTGTCATGAAGGGTCGCCTC 615

Qy        735 GCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGC 794
              || |||||||||| |||||||||||| | | |||||| |||||||||| |||||||
```

```
Db        616 GCCGACGCCAAGGCTGTGCTGGAGAAGATCGCTGACACGCCGGAGGAGGCGTCGGAGCGG 675

Qy        795 CTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACC 854
              || || |||||||||||| || || |||||||| || || |||||||||||||||||| |||
Db        676 CTCGCTGACATCAAGGCGGCGGCTGGCATCCCAGATGACCTCGACGGCGACGTGGTCACC 735

Qy        855 GTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCG 914
              || |||||| | |||| | ||| || ||| |||| ||||||| || |||||||| 
Db        736 GTGTCCAAGAAGAGAGGCGGCGAGGAGGGGCAGGTTTGGAGGGAGCTCGTCGTGTCCCCG 795

Qy        915 ACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCG 974
              |||||||||||||||| ||| |||| | | | ||||| |||||||||||||||| |||
Db        796 ACCCCGGCCATGCGGCGAATCGTGCTCGCGGCGGTCGGCCTCCACTTCTTCCAGCAGGCG 855

Qy        975 TTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACG 1034
              | |||| | |||||||||| |||||||||| | ||||||| ||||| | || | ||
Db        856 TCGGGCGTCGACTCCGTCGTGCTCTACAGCCCGCGCGTGTTCCAGAGCGCGGGCATCACC 915

Qy       1035 AACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATC 1094
                |||| | ||  || |||||||||| ||  |  |||| || |||||| | | ||||||
Db        916 GGCGACGACCAGCTCCTGGGCACCACCTGCGCCGTCGGCTTCGCCAAGACGGTCTTCATC 975

Qy       1095 TTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACG 1154
                || | ||  | ||| |  ||||| ||| ||| |||||||||||| |||||  |||||||
Db        976 CTGGTCGCCGCGTTCCTGCTCGACCGCGCCGGCCGGCGGCCGCTGCTGCTGACCAGCACG 1035

Qy       1155 GGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCAC 1214
              ||||||||  || |||||||| |||||||||  | |||||||||||||| |||      |
Db       1036 GGCGGGGATGGTCTTCTCCCTCGTCGGCCTCGCGACGGGGCTCACCGTCGTGGGCGGCAGC 1095

Qy       1215 CCCGACGCCAAGATACC---TTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTAC 1271
              || |||||| || | ||    |||||| ||||||| || |||| |||| ||| |||||||
Db       1096 CCGGACGCCCAGGTGCCCAGCTGGGCCGTCGGCCTGTGCGTCGCGTCCATCCTGGCCTAC 1155

Qy       1272 GTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTC 1331
              |||||||||||||| ||||||| || ||||| |  | |||||||| ||||||||||||||
Db       1156 GTCGCCTTCTTCTCCGTCGGCCTCGGGCCCATGTCCGGCGTGTACACCTCGGAGATCTTC 1215

Qy       1332 CCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGC 1391
              ||||| | || |||||||||||||| | || ||| ||| ||  ||||||||||||||||
Db       1216 CCGCTGCGGGCGCGCGCGCTGGGCTTCGCGGTCGCCGTGGCGTGCAACCGCGTCACCAGC 1275

Qy       1392 GGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTC 1451
              |||||||||||||||||||||||||||||||| |||| ||||||||||||||||||||||
Db       1276 GGCGTCATCTCCATGACCTTCCTGTCGCTCTCCAGCGCCATCACCATCGGCGGCAGCTTC 1335

Qy       1452 TTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAG 1511
              |||||||| ||| |||| | ||||||| |||||||||||| |||   ||||||||||
Db       1336 TTCCTCTACGCCGCCATCTCGTCGCTCGCGTGGGTGTTCTTCTTCACTCGCCTCCCGGAG 1395

Qy       1512 ACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCG 1571
              || ||||||| ||||||||||||||| ||||| |||||||      | | | |||  |
Db       1396 ACGCGCGGCCAGACGCTGGAGGAGATCGGCAAGGTGTTCGGCATGGACGACACGGCCATG 1455

Qy       1572 GAATCAGACGAGCCAGCCAAGGAGAAGAA----------------------GAAGGTG 1607
              ||| |||| ||  | |||  | | | | |                      ||||||
Db       1456 GAAGCAGAGGACTCCGCCGCGTACAGGGAAAGACTCTTGGCAACTTCTCCATGAAAGGTG 1515

Qy       1608 GAAATGGCCGCCACTAACTGATCAAACTAACCGCAAAATCACCAAATCCTAAGGGTTTTC 1667
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1516 GAAATGGCCGCCACTAACTGATCAAACTAACCGCAAAATCACCAAATCCTAAGGGTTTTC 1575

Qy       1668 TTGCAAAAACGTGTGCTGTACTGGCTAGCTAGCAAGTAGTAGCAGCAACGTGGGAAGATT 1727
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1576 TTGCAAAAACGTGTGCTGTACTGGCTAGCTAGCAAGTAGTAGCAGCAACGTGGGAAGATT 1635

Qy       1728 CGCTGATCCGGCGTTGCTGGAGAGCGACGGCCGGCGACGACAAAGCTGAGCTCCAGCTCG 1787
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1636 CGCTGATCCGGCGTTGCTGGAGAGCGACGGCCGGCGACGACAAAGCTGAGCTCCAGCTCG 1695

Qy       1788 AGACTTCTTAAAATCATCTTCAAGTACATGGATTTTATTTTGCTCTTTGCTTTGTCCGTA 1847
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1696 AGACTTCTTAAAATCATCTTCAAGTACATGGATTTTATTTTGCTCTTTGCTTTGTCCGTA 1755
```

```
Qy       1848 AAAGTTGTACTATGCGATGAAGAATACCAGTATGTAGCAAGGCTGAGGTTGTGTGTAGCT 1907
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1756 AAAGTTGTACTATGCGATGAAGAATACCAGTATGTAGCAAGGCTGAGGTTGTGTGTAGCT 1815

Qy       1908 ACTAGAAGTGTCAGTCACGTTGTTCTTGTAAGAAATGTTTAACTGTTAATTAAGCAGTAT 1967
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1816 ACTAGAAGTGTCAGTCACGTTGTTCTTGTAAGAAATGTTTAACTGTTAATTAAGCAGTAT 1875

Qy       1968 TGTTGCAGTAATCAAAAA 1985
             ||||||||||||||| ||
Db       1876 TGTTGCAGTAATCAACAA 1893
```

RESULT 4
US-10-051-902-25
; Sequence 25, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 25
; LENGTH: 2089
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-051-902-25

Query Match 54.3%; Score 1094.6; DB 14; Length 2089;
Best Local Similarity 79.1%; Pred. No. 1.1e-267;
Matches 1343; Conservative 1; Mismatches 335; Indels 18; Gaps 3;

```
Qy         84 GCACCACCGGAGATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGC 143
              ||  |  ||  |||||    |||||| || |||| ||||||| ||| ||||||||||||
Db         65 GCGTCGTCGACGATGGACCGCGCCGCACTCCCGGCGGCCGTCGAGCCCAAGAAGAAGGGC 124

Qy        144 AACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGC 203
              |||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 AACGTGAGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGC 184

Qy        204 TACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGT 263
              |||||  |||||  |||||||||||  ||||||||||||||||  |||||||   | ||  ||||
Db        185 TACGACATCGGCGTGATGAGCGGAGCGTCGCTGTACATCCAGAAGGATCTGAAGATCAAC 244

Qy        264 GACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTC 323
              |||    ||  |||||||  ||||||||||  ||  |||  |  ||||||||||  |||||||||
Db        245 GACACCCAGCTGGAGGTCCTCATGGGCATCCTCAACGTGTACTCGCTCATTGGCTCCTTC 304

Qy        324 GCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTC 383
              ||||||||||||||||  ||||||||||  |  |||||||||  ||||||||||||||||
Db        305 GCGGCGGGGCGGACGTCCGACTGGATCGGCCGGCGCTTCACCATCGTCTTCGCCGCCGTC 364

Qy        384 ATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTC 443
              ||  ||||||||||  |:  ||  |||||||  |||  ||||||||||||||||||||||||
Db        365 ATCTTCTTCGCGGGCGCCCTCATCATGGGCTTCTCCGTCAACTACGCCATGCTCATGTTC 424

Qy        444 GGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACC 503
              ||  |||||||||||||||||||||||||  |||||  ||||||||||||  |||  ||||
Db        425 GGGCGCTTCGTGGCCGGCATCGGCGTGGGGTACGCTCTCATGATCGCGCCCGTGAACACG 484

Qy        504 GCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATC 563
              |  |||||||||  ||  |||||  ||||||  ||  ||  ||  ||||||||||||||||||
Db        485 GGCGAGGTGTCCCCCGCGTCTGCCCGTGGGGTTCTCACATCCTTCCCGGAGGTGTTCATC 544

```
Qy   564 AACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAAC 623
              |||||||||||||| ||||| || ||||| |||| || ||| ||||| | || |
Db   545 AACTTCGGCATCCTCCTCGGATATGTCTCCAACTTCGCCTTCGCCCGCCTCTCCCTCCGC 604

Qy   624 CTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATG 683
              ||||| ||||||||| |||||||||| ||||||| ||||||||| ||||||||| |||||
Db   605 CTCGGCTGGCGCATTATGCTCGGCATAGGCGCGGTGCCGTCCGTCCTGCTCGCGTTCATG 664

Qy   684 GTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCC 743
              |||||||||||||| ||||| || |||||||| |||||||||| || ||||||||||||
Db   665 GTGCTCGGCATGCCCGAGTCTCCCCGGTGGCTCGTCATGAAGGGCCGTCTCGCGGACGCC 724

Qy   744 AAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGAC 803
              ||||| ||||| | ||||| ||||||||| |||| |||||||| |||||| | ||||||
Db   725 AAGGTTGTGCTTGCCAAGACGTCCGACACGCCGGAAGAGGCCGCCGAGCGCATCGCCGAC 784

Qy   804 ATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCC--- 860
              || ||| | |||||||||||||||| || |||||||||||||||||| |||| |||
Db   785 ATTAAGACTGCCGCCGGCATCCCTCTGGGCCTCGACGGCGACGTGGTCCCCGTGCCCAAA 844

Qy   861 AAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCG 920
              || | ||| ||| | | |||||||| || | |||||| |||||||||| ||||||
Db   845 AACAAAGGAAGCAGCGAGGAGAAGCGCGTTTTGAAGGACCTCATCCTGTCACCGACCATA 904

Qy   921 GCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGC 980
              |||||||| | |||||| | | || |||||||||||||||||||||| | ||||
Db   905 GCCATGCGCCACATCCTCATCGCGGGAATCGGCATCCACTTCTTCCAGCAGTCTTCGGGC 964

Qy   981 ATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGAC 1040
              || || ||||||| |||||||||| || || ||||||||| |||| | ||| ||||
Db   965 ATCGACGCCGTCGTGCTCTACAGCCCGCTAGTTTTCAAGAGCGCCGGCATCACGGGCGAC 1024

Qy   1041 AAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTG 1100
              | | || |||||||| | || |||| | ||||| | | | |||||| || ||
Db   1025 AGCCGTCTCCGCGGCACCACCGTGGCGGTCGGGGCCACCAATACGGTCTTCATCCTGGTG 1084

Qy   1101 GCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGG 1160
              || || ||| || ||||| || || | |||||||||||| ||||| ||||||||||||
Db   1085 GCCACCTTCCTCCTCGACCGCATCCGCCGGCGGCCGCTGGTGCTGACCAGCACGGGCGGC 1144

Qy   1161 ATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGAC 1220
              || || ||||| | | ||||||| | |||||||||||| || ||| ||||| |||
Db   1145 ATGCTCGTCTCCTTAGTGGGCCTCGCGACGGGGCTCACCGTCATCAGCCGCCACCCGGAC 1204

Qy   1221 GCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTC 1280
              | ||||| | |||||||||| ||| ||||| || || | | |||||||| ||||||
Db   1205 GAGAAGATCACCTGGGCCATCGTCCTGTGCATCTTCTGCATCATGGCCTACGTGGCCTTC 1264

Qy   1281 TTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAG 1340
              |||||||||||||| |||||||||||||||||||||||||||||||||||||||||| ||
Db   1265 TTCTCCATCGGCCTCGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTGCAC 1324

Qy   1341 GTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATC 1400
              |||||||||||||||||||||||||||||| || ||||| |||| ||||||| | |||||||||| |||
Db   1325 GTGCGCGCGCTGGGCTGCTCCCTGGGCGTGGCCGTCAACCGCCTGACCAGCGGCGTGATC 1384

Qy   1401 TCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTAC 1460
              |||||||||||| | |||||||||||||||||| |||||||||||| |||||||||| |
Db   1385 TCCATGACCTTCATTTCGCTGTCCAAGGCCATGACCATCGGCGGCGCCTTCTTCCTCTTC 1444

Qy   1461 TCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGC 1520
              |||||||||| | |||| |||||||||||||| | ||||||| ||||||||||||||||
Db   1445 GCCGGCATCGCCTCATTCGCATGGGTGTTCTTCTTCGCCTACCTGCCGGAGACCCGCGGC 1504

Qy   1521 CGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGAC 1580
              || |||||||||| ||||||| |||||||||| |||||||| | | | | | |
Db   1505 CGCACGCTGGAGGACATGAGCTCGCTGTTCGGCAACACGGCCACGCACAAGCAGGGCGCC 1564

Qy   1581 GAGCCAGCCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCACTAACTGATCAAACTAACCG 1640
              | | |||| | || | | || | | | || | | | | |||| ||||
Db   1565 GCGGAAGCCGACGACGACGCCGGGGAGAAGAAGGTGGAAATGGCCGCCACCAACTGACCG 1624

Qy   1641 CAAAATCACCAAATCCTAAGGGTTTTCTTGCAAAAACGTGTGCTGTACTGGCTAGCTAGC 1700
```
```

```
Db       1625 CAAGTTGGCAGATCGCGATGCGAAGACTTGCGCTGTATCCGTCTCGGCTAGCTAGCTGCC 1684

Qy       1701 A-----------AGTAGTAGCAGCAACGTGGGAAGATTCGCTGATCCGGC----GTTGCT 1745

Db       1685 ACAAGGCCACATAGATGACGAAGTAGCGTGGGAAGATTCGCTGATCCGGCCGGAGCTGCC 1744

Qy       1746 GGAGAGCGACGGCCGGC 1762

Db       1745 GGAGGGCGACGGCAAGC 1761
```

RESULT 5
US-10-051-909-25
; Sequence 25, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB1163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 25
; LENGTH: 2089
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-051-909-25

Query Match 54.3%; Score 1094.6; DB 14; Length 2089;
Best Local Similarity 79.1%; Pred. No. 1.1e-267;
Matches 1343; Conservative 1; Mismatches 335; Indels 18; Gaps 3;

```
Qy         84 GCACCACCGGAGATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGC 143

Db         65 GCGTCGTCGACGATGGACCGCGCCGCACTCCCGGCGGCCGTCGAGCCCAAGAAGAAGGGC 124

Qy        144 AACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGC 203

Db        125 AACGTGAGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGC 184

Qy        204 TACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGT 263

Db        185 TACGACATCGGCGTGATGAGCGGAGCGTCGCTGTACATCCAGAAGGATCTGAAGATCAAC 244

Qy        264 GACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTC 323

Db        245 GACACCCAGCTGGAGGTCCTCATGGGCATCCTCAACGTGTACTCGCTCATTGGCTCCTTC 304

Qy        324 GCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTC 383

Db        305 GCGGCGGGGCGGACGTCCGACTGGATCGGCCGGCGCTTCACCATCGTCTTCGCCGCCGTC 364

Qy        384 ATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTC 443

Db        365 ATCTTCTTCGCGGGCGCCCTCATCATGGGCTTCTCCGTCAACTACGCCATGCTCATGTTC 424

Qy        444 GGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACC 503

Db        425 GGGCGCTTCGTGGCCGGCATCGGCGTGGGGTACGCTCTCATGATCGCGCCCGTGAACACG 484

Qy        504 GCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATC 563

Db        485 GGCGAGGTGTCCCCCGCGTCTGCCCGTGGGGTTCTCACATCCTTCCCGGAGGTGTTCATC 544

Qy        564 AACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAAC 623

Db        545 AACTTCGGCATCCTCCTCGGATATGTCTCCAACTTCGCCTTCGCCCGCCTCTCCCTCCGC 604

Qy        624 CTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATG 683

Db        605 CTCGGCTGGCGCATTATGCTCGGCATAGGCGCGGTGCCGTCCGTCCTGCTCGCGTTCATG 664

Qy        684 GTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCC 743

Db        665 GTGCTCGGCATGCCCGAGTCTCCCCGGTGGCTCGTCATGAAGGGCCGTCTCGCGGACGCC 724

Qy        744 AAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGAC 803

Db        725 AAGGTTGTGCTTGCCAAGACGTCCGACACGCCGGAAGAGGCCGCCGAGCGCATCGCCGAC 784

Qy        804 ATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCC--- 860

Db        785 ATTAAGACTGCCGCCGGCATCCCTCTGGGCCTCGACGGCGACGTGGTCCCCGTGCCCAAA 844

Qy        861 AAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCG 920

Db        845 AACAAAGGAAGCAGCGAGGAGAAGCGCGTTTTGAAGGACCTCATCCTGTCACCGACCATA 904

Qy        921 GCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGC 980

Db        905 GCCATGCGCCACATCCTCATCGCGGGAATCGGCATCCACTTCTTCCAGCAGTCTTCGGGC 964

Qy        981 ATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGAC 1040

Db        965 ATCGACGCCGTCGTGCTCTACAGCCCGCTAGTTTTCAAGAGCGCCGGCATCACGGGCGAC 1024

Qy       1041 AAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTG 1100

Db       1025 AGCCGTCTCCGCGGCACCACCGTGGCGGTCGGGGCCACCAATACGGTCTTCATCCTGGTG 1084

Qy       1101 GCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGG 1160

Db       1085 GCCACCTTCCTCCTCGACCGCATCCGCCGGCGGCCGCTGGTGCTGACCAGCACGGGCGGC 1144

Qy       1161 ATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGAC 1220

Db       1145 ATGCTCGTCTCCTTAGTGGGCCTCGCGACGGGGCTCACCGTCATCAGCCGCCACCCGGAC 1204

Qy       1221 GCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTC 1280

Db       1205 GAGAAGATCACCTGGGCCATCGTCCTGTGCATCTTCTGCATCATGGCCTACGTGGCCTTC 1264

Qy       1281 TTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAG 1340

Db       1265 TTCTCCATCGGCCTCGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTGCAC 1324

Qy       1341 GTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATC 1400

Db       1325 GTGCGCGCGCTGGGCTGCTCCCTGGGCGTGGCCGTCAACCGCCTGACCAGCGGCGTGATC 1384

Qy       1401 TCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTAC 1460

Db       1385 TCCATGACCTTCATTTCGCTGTCCAAGGCCATGACCATCGGCGGCGCCTTCTTCCTCTTC 1444

Qy       1461 TCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGC 1520

Db       1445 GCCGGCATCGCCTCATTCGCATGGGTGTTCTTCTTCGCCTACCTGCCGGAGACCCGCGGC 1504

Qy       1521 CGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGAC 1580

Db       1505 CGCACGCTGGAGGACATGAGCTCGCTGTTCGGCAACACGGCCACGCACAAGCAGGGCGCC 1564

Qy       1581 GAGCCAGCCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCACTAACTGATCAAACTAACCG 1640

Db       1565 GCGGAAGCCGACGACGACGCCGGGGAGAAGAAGGTGGAAATGGCCGCCACCAACTGACCG 1624

Qy       1641 CAAAATCACCAAATCCTAAGGGTTTTCTTGCAAAAACGTGTGCTGTACTGGCTAGCTAGC 1700
```

```
Db       1625 CAAGTTGGCAGATCGCGATGCGAAGACTTGCGCTGTATCCGTCTCGGCTAGCTAGCTGCC 1684

Qy       1701 A-----------AGTAGTAGCAGCAACGTGGGAAGATTCGCTGATCCGGC----GTTGCT 1745
              |           ||  |  | || | ||||||||||||||||||||||||     | |||
Db       1685 ACAAGGCCACATAGATGACGAAGTAGCGTGGGAAGATTCGCTGATCCGGCCGGAGCTGCC 1744

Qy       1746 GGAGAGCGACGGCCGGC 1762
              |||| ||||||||  ||
Db       1745 GGAGGGCGACGGCAAGC 1761
```

RESULT 6
US-10-425-114-30198
; Sequence 30198, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 30198
; LENGTH: 1093
; TYPE: DNA
; ORGANISM: Oryza sativa cypress
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-OSROCYP002H08_FLI
US-10-425-114-30198

Query Match 51.6%; Score 1039.8; DB 13; Length 1093;
Best Local Similarity 96.9%; Pred. No. 6.9e-254;
Matches 1059; Conservative 1; Mismatches 33; Indels 0; Gaps 0;

```
Qy         35 CTCGACCGCCACTACTGTACACGGCCCAGAGCGAGCCTCCTCCTCCTCTGCACCACCGGA 94
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTCGACCGCCACTACTGTACACGGCCCAGAGCGAGCCTCCTCCTCCTCTGCACCACCGGA 60

Qy         95 GATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTT 154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTT 120

Qy        155 CGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGG 214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGG 180

Qy        215 GGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGT 274
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGT 240

Qy        275 GGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCG 334
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCG 300

Qy        335 GACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGC 394
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGC 360

Qy        395 GGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGT 454
              |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGGGGCGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGT 420

Qy        455 GGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTC 514
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTC 480

Qy        515 GCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCAT 574
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCAT 540

Qy        575 CCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCG 634
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCG 600

Qy        635 CATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCAT 694
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCAT 660

Qy        695 GCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCT 754
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCT 720

Qy        755 GGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGC 814
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGC 780

Qy        815 CGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGG 874
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CGCCGGCATCCCCGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGG 840

Qy        875 AAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCAT 934
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCAT 900

Qy        935 CCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGT 994
              |||||||||| || |||||||||||||||||||||||||| |||| |||||| ||||||||||
Db        901 CCTGCTGTCGGGCATCGGCATCCACTTCTTCCAGCAGGCGTCGGGCATCGACTCCGTCGT 960

Qy        995 CTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGG 1054
              |||||||||| | |||||||||||||| |||| | || |||||||||| || | ||
Db        961 GCTCTACAGCCCGCGCGTGTTCAAGAGCGCCGGCATCACCGACGACAAACACCTCCTCGG 1020

Qy       1055 CACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCAT 1114
              |||||| || | |||||||||||||| ||| |||||| | ||||||| |||||| |
Db       1021 CACCACCTGCGCCGTCGGTGTCACCAAGACGCTCTTCATCCTCGTGGCGACCTTCTTCCT 1080

Qy       1115 CGACGGCGTCGGG 1127
              |||| ||||||||
Db       1081 CGACCGCGTCGGG 1093
```

RESULT 7
US-10-051-902-19
; Sequence 19, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 19
; LENGTH: 1914
; TYPE: DNA
; ORGANISM: Zea mays
US-10-051-902-19

Query Match 49.8%; Score 1005; DB 14; Length 1914;
Best Local Similarity 79.2%; Pred. No. 6.6e-245;
Matches 1220; Conservative 1; Mismatches 311; Indels 9; Gaps 2;

```
Qy         95 GATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTT 154
              |||||||||||  || |||  |  |||||||||  ||| | ||||||||||||||||  ||
```

```
Db        55 GATGGCTTCCGACGAGCTCGCAAAGGCCGTCGAGCCCAGGAAGAAGGGCAACGTCAAGTA 114

Qy       155 CGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGG 214

Db       115 TGCCTCCATATGTGCCATCCTGGCCTCCATGGCCTCTGTCATCCTTGGCTATGACATTGG 174

Qy       215 GGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGT 274

Db       175 GGTGATGAGTGGAGCGGCCATGTACATCAAGAAGGACCTGAATATCACGGACGTGCAGCT 234

Qy       275 GGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCG 334

Db       235 GGAGATCCTGATCGGGATCCTCAGTCTCTACTCGCTGTTCGGATCCTTCGCTGGCGCGCG 294

Qy       335 GACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGC 394

Db       295 GACGTCCGACAGGATCGGGCGCCGCTTGACCGTCGTGTTCGCCGCTGTCATCTTCTTCGT 354

Qy       395 GGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGT 454

Db       355 GGGCTCGTTGCTCATGGGTTTCGCCGTCAACTACGGCATGCTCATGGCGGGCCGCTTCGT 414

Qy       455 GGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTC 514

Db       415 GGCCGGAGTCGGTGTGGGCTACGGGGGCATGATCGCGCCCGTGTACACGGCCGAGATCTC 474

Qy       515 GCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCAT 574

Db       475 GCCTGCGGCGTCCCGTGGCTTCCTGACCACCTTCCCGGAGGTGTTCATCAACATCGGCAT 534

Qy       575 CCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCG 634

Db       535 CCTGCTTGGCTACCTGTCCAACTTCGCGTTCGCGCGCCTCCCGCTCCACCTCGGCTGGCG 594

Qy       635 CATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCAT 694

Db       595 CGTCATGCTCGCCATTGGCGCAGTTCCGTCCGGCCTGCTCGCGCTCCTGGTGTTCTGCAT 654

Qy       695 GCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCT 754

Db       655 GCCCGAGTCGCCTCGGTGGCTGGTCTTGAAGGGCCGCCTCGCGGACGCCAGGGCTGTGCT 714

Qy       755 GGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGC 814

Db       715 AGAGAAGACCTCTGCCACGCCAGAGGAGGCCGCCGAGCGGCTGGCCGACATCAAGGCCGC 774

Qy       815 CGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCC---CAAGAGAGGGAG 871

Db       775 GGCGGGGATTCCGAAGGGCCTCGACGGGGACGTAGTCACCGTACCCGGCAAGGAGCAAGG 834

Qy       872 CGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCG 931

Db       835 CGGCGGTGAGTTGCAGGTGTGGAAGAAGCTCATCCTGTCCCCGACCCCGGCTGTCCGACG 894

Qy       932 CATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGT 991

Db       895 CATACTGCTCTCGGCCGTGGGTCTCCACTTCTTCCAGCAGGCTTCTGGCAGCGACTCCGT 954

Qy       992 CGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTT 1051

Db       955 CGTCCAGTACAGCGCCCGCCTGTTCAAGAGCGCGGGGATCACCGACGACAACAAGCTCCT 1014

Qy      1052 GGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTT 1111

Db      1015 GGGCGTCACCTGCGCGGTGGGCGTGACCAAGACGTTCTTCATCCTGGTGGCCACGTTCCT 1074

Qy      1112 CATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTC 1171

Db      1075 GCTGGACCGCGCGGGGCGTCGGCCTCTGCTGCTGATCAGCACGGGCGGGATGATTGTCTC 1134

Qy      1172 CCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACC 1231

Db      1135 GCTCATCTGCCTCGGGTCGGGGCTCACCGTCGCGGGGCATCACCCGGACACCAAGGTCGC 1194

Qy      1232 TTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGG 1291

Db      1195 GTGGGCCGTCGCCCTGTGCATCGCGTCAACCCTGTCCTACATCGCCTTCTTCTCCATCGG 1254

Qy      1292 CCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCT 1351

Db      1255 CCTCGGGCCCATCACGGGCGTGTACACCTCGGAAATATTCCCGCTGCAGGTGCGCGCGCT 1314

Qy      1352 GGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTT 1411

Db      1315 GGGCTTCGCGGTGGGTGTGGCGAGCAACCGCGTCACCAGCGCCGTCATCTCCATGACCTT 1374

Qy      1412 CCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGC 1471

Db      1375 CCTGTCCCTCTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGC 1434

Qy      1472 CGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGA 1531

Db      1435 CGCGGTCGCTTGGGTTTTCTTCTTCACGTGCCTCCCGGAGACACGCGGCCGGACGCTGGA 1494

Qy      1532 GGAGATGAGCAAGCTGTTCGGC------GACACGGCCGCCGCCTCGGAATCAGACGAGCC 1585

Db      1495 GGAGATGGGCAAGCTGTTCGGCATGCCAGACACGGGCATGGCTGAAGAAGCAGAAGACGC 1554

Qy      1586 AGCCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCACTAACT 1626

Db      1555 CGCAGCCAAGGAGAAGGTGGTGGAACTGCCTAGCAGCAAGT 1595
```

RESULT 8
US-10-051-909-19
; Sequence 19, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB1163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 19
; LENGTH: 1914
; TYPE: DNA
; ORGANISM: Zea mays
US-10-051-909-19

Query Match 49.8%; Score 1005; DB 14; Length 1914;
Best Local Similarity 79.2%; Pred. No. 6.6e-245;
Matches 1220; Conservative 1; Mismatches 311; Indels 9; Gaps 2;

```
Qy         95 GATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTT 154

Db         55 GATGGCTTCCGACGAGCTCGCAAAGGCCGTCGAGCCCAGGAAGAAGGGCAACGTCAAGTA 114

Qy        155 CGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGG 214

Db        115 TGCCTCCATATGTGCCATCCTGGCCTCCATGGCCTCTGTCATCCTTGGCTATGACATTGG 174

Qy        215 GGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGT 274

Db        175 GGTGATGAGTGGAGCGGCCATGTACATCAAGAAGGACCTGAATATCACGGACGTGCAGCT 234

Qy        275 GGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCG 334

Db        235 GGAGATCCTGATCGGGATCCTCAGTCTCTACTCGCTGTTCGGATCCTTCGCTGGCGCGCG 294
```

```
Qy       335 GACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGC 394
Db       295 GACGTCCGACAGGATCGGGCGCCGCTTGACCGTCGTGTTCGCCGCTGTCATCTTCTTCGT 354

Qy       395 GGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGT 454
Db       355 GGGCTCGTTGCTCATGGGTTTCGCCGTCAACTACGGCATGCTCATGGCGGGCCGCTTCGT 414

Qy       455 GGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTC 514
Db       415 GGCCGGAGTCGGTGTGGGCTACGGGGGCATGATCGCGCCCGTGTACACGGCCGAGATCTC 474

Qy       515 GCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCAT 574
Db       475 GCCTGCGGCGTCCCGTGGCTTCCTGACCACCTTCCCGGAGGTGTTCATCAACATCGGCAT 534

Qy       575 CCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCG 634
Db       535 CCTGCTTGGCTACCTGTCCAACTTCGCGTTCGCGCGCCTCCCGCTCCACCTCGGCTGGCG 594

Qy       635 CATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCAT 694
Db       595 CGTCATGCTCGCCATTGGCGCAGTTCCGTCCGGCCTGCTCGCGCTCCTGGTGTTCTGCAT 654

Qy       695 GCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCT 754
Db       655 GCCCGAGTCGCCTCGGTGGCTGGTCTTGAAGGGCCGCCTCGCGGACGCCAGGGCTGTGCT 714

Qy       755 GGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGC 814
Db       715 AGAGAAGACCTCTGCCACGCCAGAGGAGGCCGCCGAGCGGCTGGCCGACATCAAGGCCGC 774

Qy       815 CGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCC---CAAGAGAGGGAG 871
Db       775 GGCGGGGATTCCGAAGGGCCTCGACGGGGACGTAGTCACCGTACCCGGCAAGGAGCAAGG 834

Qy       872 CGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCG 931
Db       835 CGGCGGTGAGTTGCAGGTGTGGAAGAAGCTCATCCTGTCCCCGACCCCGGCTGTCCGACG 894

Qy       932 CATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGT 991
Db       895 CATACTGCTCTCGGCCGTGGGTCTCCACTTCTTCCAGCAGGCTTCTGGCAGCGACTCCGT 954

Qy       992 CGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTT 1051
Db       955 CGTCCAGTACAGCGCCCGCCTGTTCAAGAGCGCGGGGATCACCGACGACAACAAGCTCCT 1014

Qy      1052 GGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTT 1111
Db      1015 GGGCGTCACCTGCGCGGTGGGCGTGACCAAGACGTTCTTCATCCTGGTGGCCACGTTCCT 1074

Qy      1112 CATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTC 1171
Db      1075 GCTGGACCGCGCGGGGCGTCGGCCTCTGCTGCTGATCAGCACGGGCGGGATGATTGTCTC 1134

Qy      1172 CCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACC 1231
Db      1135 GCTCATCTGCCTCGGGTCGGGGCTCACCGTCGCGGGGCATCACCCGGACACCAAGGTCGC 1194

Qy      1232 TTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGG 1291
Db      1195 GTGGGCCGTCGCCCTGTGCATCGCGTCAACCCTGTCCTACATCGCCTTCTTCTCCATCGG 1254

Qy      1292 CCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCT 1351
Db      1255 CCTCGGGCCCATCACGGGCGTGTACACCTCGGAAATATTCCCGCTGCAGGTGCGCGCGCT 1314

Qy      1352 GGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTT 1411
Db      1315 GGGCTTCGCGGTGGGTGTGGCGAGCAACCGCGTCACCAGCGCCGTCATCTCCATGACCTT 1374

Qy      1412 CCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGC 1471
Db      1375 CCTGTCCCTCTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGC 1434

Qy      1472 CGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGA 1531
Db      1435 CGCGGTCGCTTGGGTTTTCTTCTTCACGTGCCTCCCGGAGACACGCGGCCGGACGCTGGA 1494

Qy      1532 GGAGATGAGCAAGCTGTTCGGC------GACACGGCCGCCGCCTCGGAATCAGACGAGCC 1585
Db      1495 GGAGATGGGCAAGCTGTTCGGCATGCCAGACACGGGCATGGCTGAAGAAGCAGAAGACGC 1554

Qy      1586 AGCCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCACTAACT 1626
Db      1555 CGCAGCCAAGGAGAAGGTGGTGGAACTGCCTAGCAGCAAGT 1595
```

RESULT 9
US-10-051-902-27
; Sequence 27, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 27
; LENGTH: 1872
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-051-902-27

Query Match 48.7%; Score 981.6; DB 14; Length 1872;
Best Local Similarity 77.4%; Pred. No. 5.8e-239;
Matches 1188; Conservative 1; Mismatches 345; Indels 0; Gaps 0;

```
Qy        99 GCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCC 158
Db       128 GCTGCGCTCCCGGAGCCGGGGGCAGTCCATCCAAGGAACAAGGGCAATTTCAAGTACGCC 187

Qy       159 TTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTG 218
Db       188 TTCACCTGCGCCCTCTGTGCTTCCATGGCCACCATCGTCCTCGGCTACGACGTTGGGGTG 247

Qy       219 ATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAG 278
Db       248 ATGAGCGGTGCGTCGCTGTACATCAAGAGGGACCTGCAGATCACGGACGTGCAGCTGGAG 307

Qy       279 GTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACG 338
Db       308 ATCATGATGGGCATCCTGAGCGTGTACGCGCTCATCGGGTCCTTCCTCGGCGCGAGGACG 367

Qy       339 TCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGG 398
Db       368 TCCGACTGGGTCGGCCGCCGCGTCACCGTCGTCTTCGCGGCCGCCATCTTCAACAACGGC 427

Qy       399 GSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCC 458
Db       428 TCCTTGCTCATGGGCTTCGCGGTCAACTACGCCATGCTCATGGTCGGGCGCTTCGTCACC 487

Qy       459 GGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCG 518
Db       488 GGAATCGGCGTGGGCTACGCCATCATGGTCGCGCCAGTGTACACGCCCGAGGTGTCCCCG 547

Qy       519 GCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTG 578
```

```
Db        548 GCGTCGGCCCGCGGCTTCCTCACGTCTTTCACCGAGGTGTTCATCAATGTGGGCATCCTC 607

Qy        579 CTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATC 638

Db        608 CTTGGCTACGTCTCCAACTACGCCTTCGCGCGCCTCCCGCTCCACCTCAGCTGGCGCGTC 667

Qy        639 ATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCG 698

Db        668 ATGCTCGGCATCGGCGCCGTCCCGTCCGCCCTGCTTGCGCTCATGGTGTTCGGCATGCCG 727

Qy        699 GAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAG 758

Db        728 GAGTCTCCTCGCTGGCTCGTCATGAAAGGCCGCCTCGCGGACGCCAGGGCCGTTCTGGCC 787

Qy        759 AAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCC 818

Db        788 AAGACCTCCGACACGCCGGAGGAGGCCGTGGAGCGCCTTGACCAGATCAAGGCTGCCGCC 847

Qy        819 GGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAAC 878

Db        848 GGCATCCCTAGGGAACTTGACGGCGACGTGGTCGTCATGCCTAAGACAAAAGGCGGCCAG 907

Qy        879 GAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTG 938

Db        908 GAGAAGCAGGTGTGGAAGGAGCTCATCTTTTCGCCGACCCCAGCCATGCGGCGCATACTG 967

Qy        939 CTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTC 998

Db        968 CTCGCGGCGCTCGGCATCCATTTCTTTCAGCAGGCGACGGGCTCCGACTCCGTCGTGCTC 1027

Qy        999 TACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGCACC 1058

Db       1028 TATAGCCCACGCGTGTTCCAGAGCGCGGGCATCACCGGCGACAACCACCTGCTCGGCGCC 1087

Qy       1059 ACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGAC 1118

Db       1088 ACATGCGCCATGGGGGTCATGAAGACGCTCTTCATCCTGGTGGCCACGTTCCAGCTCGAC 1147

Qy       1119 GGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATC 1178

Db       1148 CGCGTCGGCAGGCGGCCGCTGCTGCTGACCAGCACGGCCGGCATGCTCGCCTGTCTCATC 1207

Qy       1179 GGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCC 1238

Db       1208 GGCCTCGGGACGGGCCTCACCGTCGTGGGTCGGCACCCGGACGCCAAGGTCCCGTGGGCC 1267

Qy       1239 ATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGC 1298

Db       1268 ATCGGCCTGTGCATCGTGTCCATCTTGGCCTACGTGTCCTTCTTCTCCATCGGCCTCGGG 1327

Qy       1299 CCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGC 1358

Db       1328 CCCCTCACCAGCGTGTACACCTCGGAGGTCTTCCCACTGCGGGTGCGCGCGCTGGGCTTC 1387

Qy       1359 TCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCG 1418

Db       1388 GCGCTGGGCACGTCATGCAACCGCGTCACCAGCGCCGCGGTCTCCATGTCCTTCCTGTCC 1447

Qy       1419 CTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTC 1478

Db       1448 TTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTGTACGCCGGCATCGCGGCGATA 1507

Qy       1479 GCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATG 1538

Db       1508 GGATGGATTTTCTTCTTCACCTTCATTCCGGAGACGCGTGGCCTGCCGCTCGAGGAGATA 1567

Qy       1539 AGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAGCCAAGGAGAAG 1598

Db       1568 GGGAAGCTTTTCGGCATGACGGACACGGCCGTCGAAGCCCAAGACACCGCCACGAAAGAC 1627

Qy       1599 AAGAAGGTGGAAATGGCCGCCACTAACTGATCAA 1632
```

```
Db        1628 AAGGCGAAAGTAGGGGAGATGAACTAGTGAGCTA 1661


RESULT 10
US-10-051-909-27
; Sequence 27, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
;  APPLICANT: Allen, Steve
;  APPLICANT:  Helentjaris, Tim
;  APPLICANT:  Hitz, Bill
;  APPLICANT:  Kinney, Tony
;  APPLICANT:  Tingey, Scott
;  TITLE OF INVENTION: Plant Sugar Transport Proteins
;  FILE REFERENCE: BB1163 US CIP
;  CURRENT APPLICATION NUMBER: US/10/051,909
;  CURRENT FILING DATE:  2002-01-17
;  PRIOR APPLICATION NUMBER: 60/083,044
;  PRIOR FILING DATE: April 24, 1998
;  NUMBER OF SEQ ID NOS: 38
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 27
;   LENGTH: 1872
;   TYPE: DNA
;   ORGANISM: Triticum aestivum
US-10-051-909-27

  Query Match            48.7%;  Score 981.6;  DB 14;  Length 1872;
  Best Local Similarity  77.4%;  Pred. No. 5.8e-239;
  Matches 1188;  Conservative    1;  Mismatches  345;  Indels    0;  Gaps    0;

Qy          99 GCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCC 158

Db         128 GCTGCGCTCCCGGAGCCGGGGGCAGTCCATCCAAGGAACAAGGGCAATTTCAAGTACGCC 187

Qy         159 TTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTG 218

Db         188 TTCACCTGCGCCCTCTGTGCTTCCATGGCCACCATCGTCCTCGGCTACGACGTTGGGGTG 247

Qy         219 ATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAG 278

Db         248 ATGAGCGGTGCGTCGCTGTACATCAAGAGGGACCTGCAGATCACGGACGTGCAGCTGGAG 307

Qy         279 GTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACG 338

Db         308 ATCATGATGGGCATCCTGAGCGTGTACGCGCTCATCGGGTCCTTCCTCGGCGCGAGGACG 367

Qy         339 TCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGG 398

Db         368 TCCGACTGGGTCGGCCGCCGCGTCACCGTCGTCTTCGCGGCCGCCATCTTCAACAACGGC 427

Qy         399 GSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCC 458

Db         428 TCCTTGCTCATGGGCTTCGCGGTCAACTACGCCATGCTCATGGTCGGGCGCTTCGTCACC 487

Qy         459 GGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCG 518

Db         488 GGAATCGGCGTGGGCTACGCCATCATGGTCGCGCCAGTGTACACGCCCGAGGTGTCCCCG 547

Qy         519 GCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTG 578

Db         548 GCGTCGGCCCGCGGCTTCCTCACGTCTTTCACCGAGGTGTTCATCAATGTGGGCATCCTC 607

Qy         579 CTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATC 638

Db         608 CTTGGCTACGTCTCCAACTACGCCTTCGCGCGCCTCCCGCTCCACCTCAGCTGGCGCGTC 667

Qy         639 ATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCG 698

Db         668 ATGCTCGGCATCGGCGCCGTCCCGTCCGCCCTGCTTGCGCTCATGGTGTTCGGCATGCCG 727

Qy         699 GAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAG 758
```

```
Db        728 GAGTCTCCTCGCTGGCTCGTCATGAAAGGCCGCCTCGCGGACGCCAGGGCCGTTCTGGCC 787

Qy        759 AAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCC 818
              |||||||||||||||| |||||||||||| ||||||||| | | | ||||||||| ||||||
Db        788 AAGACCTCCGACACGCCGGAGGAGGCCGTGGAGCGCCTTGACCAGATCAAGGCTGCCGCC 847

Qy        819 GGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAAC 878
              |||||||||  ||| || ||||||||||||||||    | | || |||| |     ||||  |
Db        848 GGCATCCCTAGGGAACTTGACGGCGACGTGGTCGTCATGCCTAAGACAAAAGGCGGCCAG 907

Qy        879 GAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTG 938
              ||||||| |||||||||||||||||||| | || |||||||| |||||||||||||| |||
Db        908 GAGAAGCAGGTGTGGAAGGAGCTCATCTTTTCGCCGACCCCAGCCATGCGGCGCATACTG 967

Qy        939 CTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTC 998
              ||  | | | |||||||||| ||||| ||||| |||  ||||    |||||||||  ||
Db        968 CTCGCGGCGCTCGGCATCCATTTCTTTCAGCAGGCGACGGGCTCCGACTCCGTCGTGCTC 1027

Qy        999 TACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGCACC 1058
              || ||||| | ||||||| ||||| | ||  | ||   |||||| ||| |  | ||| ||
Db       1028 TATAGCCCACGCGTGTTCCAGAGCGCGGGCATCACCGGCGACAACCACCTGCTCGGCGCC 1087

Qy       1059 ACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGAC 1118
              || ||  |  | || ||||  |||| ||| |||||| || |||| || |||    |||||
Db       1088 ACATGCGCCATGGGGGTCATGAAGACGCTCTTCATCCTGGTGGCCACGTTCCAGCTCGAC 1147

Qy       1119 GGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCATC 1178
              |||||||  |||||||||| |||||  |||||||| ||| ||  ||  ||  |||||||
Db       1148 CGCGTCGGCAGGCGGCCGCTGCTGCTGACCAGCACGGCCGGCATGCTCGCCTGTCTCATC 1207

Qy       1179 GGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGCC 1238
              ||||||||  | || |||||||||| || | |||||| |||||||| | || ||||||
Db       1208 GGCCTCGGGACGGGCCTCACCGTCGTGGGTCGGCACCCGGACGCCAAGGTCCCGTGGGCC 1267

Qy       1239 ATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGC 1298
              ||||||||  |||| | | ||||||||  |||||||||||||||||| ||  |||||||
Db       1268 ATCGGCCTGTGCATCGTGTCCATCTTGGCCTACGTGTCCTTCTTCTCCATCGGCCTCGGG 1327

Qy       1299 CCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGC 1358
              ||| ||||  | ||||||| ||||||| |||||| || | ||||||||||||||||| |
Db       1328 CCCCTCACCAGCGTGTACACCTCGGAGGTCTTCCCACTGCGGGTGCGCGCGCTGGGCTTC 1387

Qy       1359 TCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCG 1418
              ||||  |||    |  |||||||||||||||| ||   |||||||| ||||||||||
Db       1388 GCGCTGGGCACGTCATGCAACCGCGTCACCAGCGCCGCGGTCTCCATGTCCTTCCTGTCC 1447

Qy       1419 CTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTC 1478
              |||||||||||||||||||||||||||||||||||||| ||| ||||||||| ||| |
Db       1448 TTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTGTACGCCGGCATCGCGGCGATA 1507

Qy       1479 GCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATG 1538
              |  ||| | ||||||| ||.|| | | |||||||| || |||| | |||| ||||||||
Db       1508 GGATGGATTTTCTTCTTCACCTTCATTCCGGAGACGCGTGGCCTGCCGCTCGAGGAGATA 1567

Qy       1539 AGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAGCCAAGGAGAAG 1598
              | ||||| ||||||  |||| | | |||  ||| |  | ||  | |||| | |  |
Db       1568 GGGAAGCTTTTCGGCATGACGGACACGGCCGTCGAAGCCCAAGACACCGCCACGAAAGAC 1627

Qy       1599 AAGAAGGTGGAAATGGCCGCCACTAACTGATCAA 1632
              ||| |  | | ||     |  | ||| | |
Db       1628 AAGGCGAAAGTAGGGGAGATGAACTAGTGAGCTA 1661


RESULT 11
US-10-425-114-28112
; Sequence 28112, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 28112
;   LENGTH: 1704
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB4746-010-A4_FLI
US-10-425-114-28112

  Query Match           38.7%;  Score 779.6;  DB 13;  Length 1704;
  Best Local Similarity   71.4%;  Pred. No. 1.3e-187;
  Matches 1044;  Conservative    1;  Mismatches  405;  Indels   12;  Gaps    1;

Qy        152 GTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATAT 211
              || |||| |||||||||||| ||||||| ||||||| |||||||||||| |||||||||
Db        136 GTACGCCCTCGCCTGCGCCGTCCTCGCTTCCATGAACTCCATCCTCCTTGGCTACGATGT 195

Qy        212 CGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAA 271
              |||||||||||| ||| ||| | ||| ||| ||||| |||| |||| ||| | |
Db        196 GTCGGTGATGAGCGGAGCGCAGCTATTCATGAAGCAGGACCTCAAGATCACGGACACGCA 255

Qy        272 GGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGG 331
              | | ||| | |||   |||||  | ||  |||||||||| |||||||  ||||||||||
Db        256 GATCGAGATCCTCGCCGGCATCATCAATATCTACTCGCTCGTCGGCTCGCTCGCGGCGGG 315

Qy        332 GCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTT 391
              |||||||| |||||| |||||||||||||||| ||| |||| || | ||| |||||
Db        316 CCGGACGTCCGACTGGCTCGGCCGGCGGTACACCATGGTGCTCGCGGCGGCCATCTTCTT 375

Qy        392 CGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTT 451
              |||||| |:| || |||||||  |||||    | |||||| || |||||| | |||||||||
Db        376 CGCGGGCGCGCTCATCATGGGCCTCGCCCCGAGCTACGCGATCCTCATGCTTGGCCGCTT 435

Qy        452 CGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGT 511
              |||||| ||| |||||||| || |||||||||||||||||| || |||||||| |||||||
Db        436 CGTGGCGGGCGTCGGCGTCGGATACGCGCTCATGATCGCACCCGTGTACACGGCCGAGGT 495

Qy        512 GTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGG 571
              |||||  |||||||||| ||| | || ||||| ||||||||||||||||||||    ||
Db        496 CTCGCCCACGTCGGCGCGCGGCCTGCTCACGTCCTTCCCGGAGGTGTTCATCAACACGGG 555

Qy        572 CATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTG 631
              | || |||||||| |||||  |||| || |||  | || | ||  || |||||  | ||
Db        556 GGTGCTCCTCGGGTATGTCTCCAACTACGCCTTCCACGGCCTCCCCGTGCACCTCAGCTG 615

Qy        632 GCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGG 691
              |||| ||||| ||  | ||||||| |  ||| |  |  | |||||  ||  ||| ||||
Db        616 GCGCGTCATGTTCCTCGTCGGCGCCGTCCCGCCTATCTTCCTCGCACTGGGGGTCCTCGC 675

Qy        692 CATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGT 751
              |||||||||||||||||||||||| |||||| |||| ||| |||  |||||        ||
Db        676 CATGCCGGAGTCGCCGCGGTGGCTCGTCATGCAGGGGCGCATCGGCGACGCGCGCCGCGT 735

Qy        752 GCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGC 811
              |||||  |||||||||  || |  | |  |||||||  |||||||  || ||||||||||
Db        736 GCTGGCCAAGACCTCCAACTCCCCCGCCGAGGCCGAGGAGCGGCTCGCCGACATCAAGAA 795

Qy        812 CGCCGCCGGCATCCCTGAGG----------AGCTCGACGGCGACGTGGTGACCGTCCC 859
              |||  |||| |||| || |          |   ||||| ||  ||||||  || ||
Db        796 TGCCATCGGCGTCCCAGACGGCGTCGGCGACAACGACGACGACGTGGTGGTGGTCGCCCG 855

Qy        860 CAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCC 919
              || ||    | || |  | |   | ||||||||| ||| ||| || |  ||||||  ||
Db        856 CAGGAACAAGGGCAGCCATGGCGAAGGGGTGTGGAGGGACCTCCTCATCCGCCCGACGCC 915

Qy        920 GGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGG 979
              | || | || ||||| ||  |  || |  ||||| |||| |||||||||| || |  ||
```

```
Db        916 GCCCGTCCGCCGCATACTCATCGCCTGCCTCGGCCTCCAGTTCTTCCAGCAGGCCTCCGG 975

Qy        980 CATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGA 1039

Db        976 CATCGACTCCGTGGTGCTGTACAGCCCGCGGGTGTTCGAGAAGGCTGGGCTTAGGTCGAA 1035

Qy       1040 CAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTT 1099

Db       1036 CAACAACTCTTTGGGCGCCACCATGGCGGTGGGCGCGACCAAGACGCTGTTCATCCTGGT 1095

Qy       1100 GGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGG 1159

Db       1096 GGCCACGTTCTTCCTCGACCGCGTCGGGCGGAGACCGCTGCTGCTCACAAGCGCGGGCGG 1155

Qy       1160 GATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGA 1219

Db       1156 GATGGTGGTGTCGCTGGTGACGCTGGCCTCGGCGTTGCGCGCCATCGACCGGCTCCCAGA 1215

Qy       1220 CGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTT 1279

Db       1216 GGGACAGGCGACGTCGCTGGCGGGCGTGAGCATCGCGGCGGTTCTCACGTTCGTGGCGTC 1275

Qy       1280 CTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCA 1339

Db       1276 CTTCTCCATCGGCATGGGCCCGATCGCGTGGGTGTACAGCTCGGAGATATTCCCACTGCG 1335

Qy       1340 GGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCAT 1399

Db       1336 GCTCCGCGCGCAGGGCTGCGCACTCGGCACGGCGATGAACCGGGTCATGAGCGGCACCAT 1395

Qy       1400 CTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTA 1459

Db       1396 CACCATGTCCTTCATCTCGCTCTACAAGGCCATCACCTTCGCCGGGAGCTTCTACTTGTA 1455

Qy       1460 CTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGG 1519

Db       1456 TGCAGGCATTGCTGCTGCCGGGTGGCTGTTCATGTTCTTCTTCCTGCCGGAGACGAGGGG 1515

Qy       1520 CCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGA 1579

Db       1516 CAGCAACCTAGAGGACACCGAGAAGCTCTTCGGCGGTGGCGACCACGACGAGGACAAGGA 1575

Qy       1580 CGAGCCAGCCAAGGAGAAGAAG 1601

Db       1576 AGATGGACACGATGCGCAGAAG 1597
```

RESULT 12
US-10-437-963-78515
; Sequence 78515, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 78515
; LENGTH: 1809
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_78314C.1
US-10-437-963-78515

Query Match 37.6%; Score 759.4; DB 17; Length 1809;
Best Local Similarity 73.0%; Pred. No. 1.9e-182;
Matches 992; Conservative 1; Mismatches 357; Indels 9; Gaps 1;

```
Qy        207 GATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGAC 266

Db        220 GACATCTCGGTGATGAGCGGCGCGCAGAAGTTCATGAAGAAGGACCTGAACATCTCGGAC 279

Qy        267 GGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCG 326

Db        280 GCGAAGGTGGAGGTCCTCGCCGGCATCATCAACATCTACTCGCTCGTCGGCTCGCTCGCG 339

Qy        327 GCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATA 386

Db        340 GCGGGCCGGACGTCCGACTGGATCGGCCGTCGCTACACCATGGTGCTCGCGTCGGTCATA 399

Qy        387 TTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGC 446

Db        400 TTCTTCGCCGGCGCGCTCATCATGGGCCTCGCCCCGAGCTACGCGATCGTCATGCTCGGC 459

Qy        447 CGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCC 506

Db        460 CGCTTCGTGGCCGGCGTGGGTGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACGGCC 519

Qy        507 GAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAAC 566

Db        520 GAGGTGGCGCCGACGTCCGCCCGCGGCTTGCTCACGTCGTTCCCCGAGGTGTTCATCAAC 579

Qy        567 TTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTC 626

Db        580 GCCGGCGTCCTCCTCGGCTACGTCTCCAACTTCGCCTTCTACCGCCTCCCGCTGCACATT 639

Qy        627 GGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTG 686

Db        640 GGCTGGCGCGTCATGTTCCTAGTCGGCGCCGTGCCGCCCGCCTTCCTCGCCGTGGGCGTC 699

Qy        687 CTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAG 746

Db        700 CTCGCCATGCCGGAGTCGCCGCGGTGGCTCGTCATGCAGGGGCGCATCGGCGACGCGCGA 759

Qy        747 GTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATC 806

Db        760 CGTGTCCTCGAGAAGACCTCCGACTCGCCGGCCGAGGCCGAGGAGCGGCTCGCGGACATC 819

Qy        807 AAGGCCGCCGCCGGCATCCCTGAGGAGCTC---------GACGGCGACGTGGTGACCGTC 857

Db        820 AAGAACGCGGTCGGCATCCCCGAAGGGATCTCCGACGAGGACGAAGTTGTAGCCGTCGTC 879

Qy        858 CCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACC 917

Db        880 CACAAGAGCAGGGGCTCCCACGGCGAGGGTGTCTGGAGGGACCTGCTGCTCCGCCCGACG 939

Qy        918 CCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTG 977

Db        940 CCGGCCGTGCGCCGCATACTCATCGCCTGCCTCGGCCTCCAGTTCTTCCAGCAAGCCTCC 999

Qy        978 GGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAAC 1037

Db       1000 GGCATCGACGCCGTCGTGCTGTACAGCCCGCGGGTGTTCGACAACGCGGGCCTCCACTCC 1059

Qy       1038 GACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTG 1097

Db       1060 GACTCCGACTCCATCGGAGCCTCCGTGGCGGTGGGCGCCAGCAAGACGCTCTTCATCCTC 1119

Qy       1098 TTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGC 1157

Db       1120 GTGGCCACGTTCCTCCTCGACCGCGTCGGCCGGAGGCCGCTGCTCCTCACCAGCGCCGGC 1179

Qy       1158 GGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCC 1217

Db       1180 GGGATGGTGATCTCGCTCGTCACGCTGGCCTCGGCGCTGCACATGATCGAGCACCGCCCG 1239
```

```
Qy       1218 GACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCC 1277
Db       1240 GAGGGCCAGGCGACGGCGCTGGTGGGTCTGAGCATCGCGATGGTGCTGGTGTTCGTGGCG 1299

Qy       1278 TTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTC 1337
Db       1300 TCCTTCTCCATCGGGATGGGCCCGATCGCGTGGGTGTACAGCTCGGAGATCTTCCCGCTG 1359

Qy       1338 CAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTC 1397
Db       1360 CGGCTGCGCGCGCAGGGATGCGCGCTCGGCACGGCGATGAACCGGGTCGTGAGCGGCGCC 1419

Qy       1398 ATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTC 1457
Db       1420 GTCAGCATGTCCTTCATCTCGCTCTACAAGGCCATCACCTTCGCCGGGAGCTTCTACCTC 1479

Qy       1458 TACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGC 1517
Db       1480 TACGCCGGCATAGCTGCGGCCGGGTGGGTGTTCATGTTCTTCTTCCTGCCGGAGACGCAA 1539

Qy       1518 GGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGAC 1556
Db       1540 GGCAGGAGCCTGGAGGACACCGTCAAACTCTTCGGCGGC 1578
```

RESULT 13
US-10-425-114-30135
; Sequence 30135, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 30135
; LENGTH: 2203
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-OSFLCYP011G10_FLI
US-10-425-114-30135

Query Match 36.7%; Score 739.6; DB 13; Length 2203;
Best Local Similarity 69.1%; Pred. No. 2.3e-177;
Matches 1028; Conservative 1; Mismatches 450; Indels 9; Gaps 1;

```
Qy        104 CGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCCTTCGC 163
Db        218 CGCCGTCCCGGTGGCCGTGGCGCCGGCGAAGCGCCCGCCCATCAACAAGTACGCCTTCGG 277

Qy        164 CTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTGATGAG 223
Db        278 CTGCGCTCTCCTCGCCTCCATGAACTCCGTCCTCCTCGGCTATGATATCTCGGTGATGAG 337

Qy        224 CGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCT 283
Db        338 CGGCGCGCAGATATTCATGAAGGAGGATTTGAAGATCACCGACACGCAGATCGAGATCCT 397

Qy        284 CATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGA 343
Db        398 CGCCGGCGTCATCAACATCTACTCGCTGTTCGGCTCGCTGGCGGCGGGCATGACGTCCGA 457

Qy        344 CTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGSGTT 403
Db        458 CTGGCTGGGCCGGCGGTACACCATGGTGCTCGCCGCCGCCATCTTCTTCACGGGCGCGCT 517

Qy        404 CCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCAT 463
Db        518 CCTCATGGGCCTCGCCCCGAACTACGCGTTCCTCATGGCCGGCCGCTTCGTGGCCGGCAT 577

Qy        464 CGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTC 523
Db        578 CGGCGTCGGCTACGCGCTCATGATCGCGCCCGTGTACACCGCCGAGGTGGCGCCGACGTC 637

Qy        524 GGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGG 583
Db        638 GGCGCGCGGCTTCCTCACCTCGTTCCCGGAGGTGTTCAACAACAGCGGCATCCTCCTCGG 697

Qy        584 GTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCT 643
Db        698 CTACGTCTCCAACTTCGCCTTCGCGCGCCTCCCCGTCCACCTGAGCTGGCGCGCCATGTT 757

Qy        644 CGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTC 703
Db        758 CCTCGTCGGCGCCGTGCCGCCCATCTTCCTCGGCATCGCCGTCCTCGCCATGCCGGAGTC 817

Qy        704 GCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGAC 763
Db        818 GCCCCGGTGGCTCGTCATGCGGGGCCGCATCGAGGACGCGCGCCGCGTGCTCCTCAAGAC 877

Qy        764 CTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCAT 823
Db        878 CTCCGACTCGCCCGACGAGGCCGAGGACCGGCTCCTCGACATCAAGAAAGCCGTCGGCAT 937

Qy        824 CCCTGAG---------GAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGG 874
Db        938 CCCCGAGGACGCGTCCGATGGCGAGGACGTGGTCGCCATCGTCCGCGCGAACAAGGCCTC 997

Qy        875 AAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCAT 934
Db        998 ACAGGGAGAGGGGGTGTGGAAGGAGCTGCTGCTCAACCCGACGAGGCCCGTCCGCCGCAT 1057

Qy        935 CCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGT 994
Db       1058 GCTCGTCGCCGGGCTCGGCCTGATGTTCATCCAGCAAGCCACCGGCGTCGACTGCGTCGT 1117

Qy        995 CTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGG 1054
Db       1118 GATGTACAGCCCGCGGGTGTTCGAGCGCGCCGGCATTAAGTCGAAGACGAACTCGCTGGG 1177

Qy       1055 CACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCAT 1114
Db       1178 GGCGTCCATGGCGGTGGGCGTTTGCAAGACCTTCTTCATCCCGATCGCGACGCTCCTCCT 1237

Qy       1115 CGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCT 1174
Db       1238 CGACCGCGTCGGACGGAGGCCGCTCCTCCTCGCCAGCGGCGGCGGCATGGCCATCTTCCT 1297

Qy       1175 CATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTG 1234
Db       1298 CTTCACCCTGGCCACGTCCCTGCTCATGATGGACCGGCGGCCGGAGGGCGAGGCGAAGGC 1357

Qy       1235 GGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCT 1294
Db       1358 GCTAGGCGCCATAAGCATCGCCGCGATGCTGTCGTTCGTGGCGTCGTTCGCCTCGGGGCT 1417

Qy       1295 TGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGG 1354
Db       1418 GGGCCCCGTCGCGTGGGTGTACACGTCGGAGATCTACCCGGTGCGGCTGCGTGCGCAGGC 1477

Qy       1355 CTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCT 1414
Db       1478 GGCGGCCATCGGGACGGGGCTGAACCGGCTGATGAGCGGTGCCACCACCATGTCCTTCCT 1537

Qy       1415 GTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGC 1474
Db       1538 CTCGCTCTCCAACGCCATCACCATCGCCGGGAGCTTCTACCTGTACGCGTCCATCGCCGC 1597
```

```
Qy       1475 GCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGA 1534
             |  |||||||||  | |||  || ||| |||||||||    |||  ||  ||||||||
Db       1598 CGCAGGGTGGGTGTTCATGTACTTCTTCCTGCCGGAGACCAAGGGCAAGAGCCTGGAGGA 1657

Qy       1535 GATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGACGA 1582
              |     ||||| ||||||  |     | | || |    ||     | |||
Db       1658 CACCGTGAAGCTCTTCGGCAAGGACACAGACGACGACGATGACGTCGA 1705


RESULT 14
US-10-425-114-26589
; Sequence 26589, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 26589
;   LENGTH: 2231
;   TYPE: DNA
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB4370-026-E2_FLI
US-10-425-114-26589

  Query Match           36.6%;  Score 738;  DB 13;  Length 2231;
  Best Local Similarity  69.0%;  Pred. No. 5.9e-177;
  Matches 1027;  Conservative    1;  Mismatches  451;  Indels    9;  Gaps    1;

Qy        104 CGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCCTTCGC 163
              |||||  | |  ||  |  |   || |||||      || |  |  || |||||||||
Db        218 CGCCGTCCCGGTGGCCGTGGCGCCGGCGAAGCGCCCGCCCATCAACAAGTACGCCTTCGG 277

Qy        164 CTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTGATGAG 223
              ||||||  |||||||||||||||| |||| |||||||||||||| ||||||  |||||||||
Db        278 CTGCGCTCTCCTCGCCTCCATGAACTCCGTCCTCCTCGGCTATGATATCTCGGTGATGAG 337

Qy        224 CGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAGGTTCT 283
              |||  |||   | | | |||  ||| |||| || || ||||  |||  | || | |||  | ||
Db        338 CGGCGCGCAGATATTCATGAAGGAGGATTTGAAGATCACCGACACGCAGATCGAGATCCT 397

Qy        284 CATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTCGGA 343
              |   |||  |  | |||  |||||||||  |||||||  | ||||||||   ||||||  ||
Db        398 CGCCGGCGTCATCAACATCTACTCGCTGTTCGGCTCGCTGGCGGCGGGCATGACGTCCGA 457

Qy        344 CTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGSGTT 403
              ||||  | |||||||||||||||||  |||  |||||||| |||  ||||||  ||||  :| |
Db        458 CTGGCTGGGCCGGCGGTACACCATGGTGCTCGCCGCCGCCATCTTCTTCACGGGCGCGCT 517

Qy        404 CCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGGCAT 463
              |||||||||  |||||  ||||||||  | ||||||  |||||||||||||||||||||||
Db        518 CCTCATGGGCCTCGCCCCGAACTACGCGTTCCTCATGGCCGGCCGCTTCGTGGCCGGCAT 577

Qy        464 CGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGCGTC 523
              ||||||  |||||||||||||||||||||||||||||||||||||||||||||  ||||
Db        578 CGGCGTCGGCTACGCGCTCATGATCGCGCCCGTGTACACCGCCGAGGTGGCGCCGACGTC 637

Qy        524 GGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCTCGG 583
              ||||||  ||||||||  ||  |||||||||||||||||  ||||  |||||||||  |||||
Db        638 GGCGCGCGGCTTCCTCACCTCGTTCCCGGAGGTGTTCAACAACAGCGGCATCCTCCTCGG 697

Qy        584 GTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCATGCT 643
              ||||||||  ||||  ||  |||  |  |||  |  ||  |  ||||  |  |  ||||  ||||  |
Db        698 CTACGTCTCCAACTTCGCCTTCGCGCGCCTCCCCGTCCACCTGAGCTAGCGCGCCATGTT 757

Qy        644 CGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGAGTC 703
              |  | |||||||  |  ||||  ||  |  |  ||||   ||   ||  ||||  ||||||||||||
Db        758 CCTCGTCGGCGCCGTGCCGCCCATCTTCCTCGGCATCGCCGTCCTCGCCATGCCGGAGTC 817

Qy        704 GCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAAGAC 763
              |||  ||||||||  ||||||  |||  |||  |||  ||||||         |||||     |||||
Db        818 GCCCCGGTGGCTCGTCATGCGGGGCCGCATCGAGGACGCGCGCCGCGTGCTCCTCAAGAC 877

Qy        764 CTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGGCAT 823
              |||||||  ||  ||  |||||||  |||  ||  ||   ||||||||||    ||||  ||||||
Db        878 CTCCGACTCGCCCGACGAGGCCGAGGACCGGCTCCTCGACATCAAGAAAGCCGTCGGCAT 937

Qy        824 CCCTGAG---------GAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGG 874
              |||  |||          ||   |||  |  ||  ||  |  |||||  |   ||     |  |
Db        938 CCCCGAGGACGCGTCCGATGGCGAGGACGTGGTCGCCATCGTCCGCGCGAACAAGGCCTC 997

Qy        875 AAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCAT 934
              |  |  ||  ||||||||||||||||  |  ||   ||||||   |  ||  |  ||  |||||
Db        998 ACAGGGAGAGGGGGTGTGGAAGGAGCTGCTGCTCAACCCGACGAGGCCCGTCCGCCGCAT 1057

Qy        935 CCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGT 994
              ||  |  |||||  |||||  |    |||  |||||||  ||    |||  |  |||  ||||||
Db       1058 GCTCGTCGCCGGGCTCGGCCTGATGTTCATCCAGCAAGCCACCGGCGTCGACTGCGTCGT 1117

Qy        995 CTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGG 1054
              |  |||||||||  |  ||||||  ||  ||  ||||  |  |  |    |  |  |||   ||||
Db       1118 GATGTACAGCCCGCGGGTGTTCGAGCGCGCCGGCATTAAGTCGAAGACGAACTCGCTGGG 1177

Qy       1055 CACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCAT 1114
              |  |   |  ||  |  ||  ||   |||||   |  ||||||   |  |  |||||   ||  ||  |
Db       1178 GGCGTCCATGGCGGTGGGCGTTTGCAAGACCTTCTTCATCCCGATCGCGACGCTCCTCCT 1237

Qy       1115 CGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCT 1174
              ||||  |||||||  |||  |||||||   |  ||  |  ||||   |||||  ||    |  |||  |||
Db       1238 CGACCGCGTCGGACGGAGGCCGCTCCTCCTCGCCAGCGGCGGCGGCATGGCCATCTTCCT 1297

Qy       1175 CATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTG 1234
              |  |||  |||  |  |    ||   |  |  |  |  ||  ||  ||  ||  |  |  ||
Db       1298 CTTCACCCTGGCCACGTCCCTGCTCATGATGGACCGGCGGCCGGAGGGCGAGGCGAAGGC 1357

Qy       1235 GGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCT 1294
              |      ||  |  ||||||||||||  |  |   ||   |  |  |||  ||  |   |||  ||    ||  ||
Db       1358 GCTAGGCGCCATAAGCATCGCCGCGATGCTGTCGTTCGTGGCGTCGTTCGCCTCGGGGCT 1417

Qy       1295 TGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGG 1354
              ||||||  ||  ||||||||||||   |||||||||  ||||  |  |  |  ||||  ||||  ||
Db       1418 GGGCCCCGTCGCGTGGGTGTACACGTCGGAGATCTACCCGGTGCGGCTGCGTGCGCAGGC 1477

Qy       1355 CTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCT 1414
                    |  ||||     |  |||||  |  |  |||||  |  ||  |  |||||  |||||||
Db       1478 GGCGGCCATCGGGACGGGGCTGAACCGGCTGATGAGCGGTGCCACCACCATGTCCTTCCT 1537

Qy       1415 GTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGC 1474
              |||||  |||||  ||||||||||||  |||  |||||||  |||  |||  |    |||||||||
Db       1538 CTCGCTCTCCAACGCCATCACCATCGCCGGGAGCTTCTACCTGTACGCGTCCATCGCCGC 1597

Qy       1475 GCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGA 1534
                  |  ||||||||||  |  |||  ||  |||  ||||||||||    |||  ||  |||||||||
Db       1598 CGCAGGGTGGGTGTTCATGTACTTCTTCCTGCCGGAGACCAAGGGCAAGAGCCTGGAGGA 1657

Qy       1535 GATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGACGA 1582
                |     |||||  ||||||  |       |  |  ||  |    ||     | |||
Db       1658 CACCGTGAAGCTCTTCGGCAAGGACACAGACGACGACGATGACGTCGA 1705


RESULT 15
US-10-425-114-28469
; Sequence 28469, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
```

; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 28469
; LENGTH: 1777
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4757-033-C12_FLI
US-10-425-114-28469

Query Match 36.4%; Score 734.8; DB 13; Length 1777;
Best Local Similarity 69.1%; Pred. No. 3.4e-176;
Matches 1022; Conservative 1; Mismatches 443; Indels 12; Gaps 1;

```
Qy        136 AGAAGGGCAACGTCCGGTTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCC 195
              | ||   ||    || ||| ||| | |||     || |       |     ||||   |
Db        238 ACAATTCCATGCACCCGTTGGCCGTTGCCATATGCAGCACGCGTACGTGCGCCTCATGCT 297

Qy        196 TCCTCGGCTACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCA 255
              ||      ||| |  |||||||||||| |||  ||| | ||| ||| ||||| |||
Db        298 ATCTGTATGCAGATGTGTCGGTGATGAGCGGAGCGCAGCTATTCATGAAGCAGGACCTCA 357

Qy        256 ACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCG 315
              | |||| |||  | || | ||| | |||   |||||  | ||  |||||||||||| |||
Db        358 AGATCACGGACACGCAGATCGAGATCCTCGCCGGCATCATCAATATCTACTCGCTCGTCG 417

Qy        316 GCTCCTTCGCGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCG 375
              ||||  |||||||||| |||||||| |||||| |||||||||||||||||| ||| |||
Db        418 GCTCGCTCGCGGCGGGCCGGACGTCCGACTGGCTCGGCCGGCGGTACACCATGGTGCTCG 477

Qy        376 CCGCCGTCATATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGC 435
              | || | ||| |||||||||||  |:| || |||||||  |||||    | |||| | || |
Db        478 CGGCGGCCATCTTCTTCGCGGGCGCGCTCATCATGGGCCTCGCCCCGAGCTACACGATCC 537

Qy        436 TCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGG 495
              |||||  | |||||||||||||| ||| ||||||| || |||||||||||||||||| || |
Db        538 TCATGCTTGGCCGCTTCGTGGCGGGCGTCGGCGTCGGATACGCGCTCATGATCGCACCCG 597

Qy        496 TGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGG 555
              |||||||  |||||||| |||||  |||||||||| ||| | || |||||  |||||||||
Db        598 TGTACACGGCCGAGGTCTCGCCCACGTCGGCGCGCGGCCTGCTCACGTCCTTCCCGGAGG 657

Qy        556 TGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGC 615
              ||||||||||    ||   | || |||||||| ||||| ||||| || |||   | || | |
Db        658 TGTTCATCAACACGGGGGTGCTCCTCGGGTATGTCTCCAACTACGCCTTCCACGGCCTCC 717

Qy        616 CGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCG 675
              |  || ||||| | |||||| ||||| ||  | |||||||  |  |||  |   |  ||||
Db        718 CCGTGCACCTCAGCTGGCGCGTCATGTTCCTCGTCGGCGCCGTCCCGCCTATCTTCCTCG 777

Qy        676 CGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCG 735
              | ||   ||| |||| |||||||||||||||||||||||| |||||| |||| ||| |||
Db        778 CACTGGGGGTCCTCGCCATGCCGGAGTCGCCGCGGTGGCTCGTCATGCAGGGGCGCATCG 837

Qy        736 CGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCC 795
                |||||       |||||||  |||||||| || |  | |  |||||| ||||||  |
Db        838 GCGACGCGCGCCGCGTGCTGGCCAAGACCTCCAACTCCCCCGCCGAGGCCGAGGAGCGGC 897

Qy        796 TGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGG----------AGCTCGACGGCG 843
              | ||||||||||||   |||  |||| |||| || |          |   ||||| ||
Db        898 TCGCCGACATCAAGAATGCCATCGGCGTCCCAGACGGCGTCGGCGACAACGACGACGACG 957

Qy        844 ACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCA 903
              ||||||  || || || ||    | || |  | |   |  |||||||| ||| |||
Db        958 TGGTGGTGGTCGCCCGCAGGAACAAGGGCAGCCATGGCGAAGGGGTGTGGAGGGACCTCC 1017
```

```
Qy         904 TCCTGTCCCCGACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCT 963
               || |  |||||| ||| || | || ||||| ||  |  ||  | |||| |||| ||||
Db        1018 TCATCCGCCCGACGCCGCCCGTCCGCCGCATACTCATCGCCTGCCTCGGCCTCCAGTTCT 1077

Qy         964 TCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCC 1023
               |||||||  || |  |||||  |||||||  ||  |  ||||||||  |  ||||||  |||
Db        1078 TCCAGCAGGCCTCCGGCATCGACTCCGTGGTGCTGTACAGCCCGCGGGTGTTCGAGAAGG 1137

Qy        1024 CCGGATTAACGAACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGA 1083
               | ||  | | |    ||||  |||  |||||| ||||   | ||  | ||  | |||||||
Db        1138 CTGGGCTTAGGTCGAACAACAACTCTTTGGGCGCCACCATGGCGGTGGGCGCGACCAAGA 1197

Qy        1084 GGCTTTTCATCTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGC 1143
                |||  ||||||  ||  ||||  ||  ||||||  |||||  |||||||||||  | ||||||  |||
Db        1198 CGCTGTTCATCCTGGTGGCCACGTTCTTCCTCGACCGCGTCGGGCGGAGACCGCTGCTGC 1257

Qy        1144 TGGGCAGCACGGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCG 1203
               |     |||  |||||||||||  |  | ||  ||  |    ||  |  |  |  |  ||  |
Db        1258 TCACAAGCGCGGGCGGGATGGTGGTGTCGCTGGTGACGCTGGCCTCGGCGTTGCGCGCCA 1317

Qy        1204 TCGGCCAGCACCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCC 1263
               ||| ||  ||  ||| ||  |   ||     |  |  |     ||| |  |||||||  |    |
Db        1318 TCGACCGGCTCCCAGAGGGACAGGCGACGTCGCTGGCGGGCGTGAGCATCGCGGCGGTTC 1377

Qy        1264 TCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGG 1323
               ||  |  |||  ||  |  ||||||||||||  |  |||||  |||  |||||||||||||||||
Db        1378 TCACGTTCGTGGCGTCCTTCTCCATCGGCATGGGCCCGATCGCGTGGGTGTACAGCTCGG 1437

Qy        1324 AGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCG 1383
               ||||  |||||  ||  |  |  |  |||||||  |||||||  |  ||||||    ||      |||||  |
Db        1438 AGATATTCCCACTGCGGCTCCGCGCGCAGGGCTGCGCACTCGGCACGGCGATGAACCGGG 1497

Qy        1384 TCACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCG 1443
               |||  ||||||  ||||  |||||  |||||  |  |||||  |  ||||||||||||  |||  ||
Db        1498 TCATGAGCGGCACCATCACCATGTCCTTCATCTCGCTCTACAAGGCCATCACCTTCGCCG 1557

Qy        1444 GCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACC 1503
               | |||||||  |  |  ||   | |||||  ||  ||    ||  |||  |||||  |  |  |   ||  ||
Db        1558 GGAGCTTCTACTTGTATGCAGGCATTGCTGCTGCCGGGTGGCTGTTCATGTTCTTCTTCC 1617

Qy        1504 TCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGGCCG 1563
               |  |||||||||  |  |||  |  |  ||  |||||  |       |||||  |||||||       |  |
Db        1618 TGCCGGAGACGAGGGGCAGCAACCTAGAGGACACCGAGAAGCTCTTCGGCGGTGGCGACC 1677

Qy        1564 CCGCCTCGGAATCAGACGAGCCAGCCAAGGAGAAGAAG 1601
                  ||  |  |||     ||  ||    |   |  |  |  |  |||||
Db        1678 ACGACGAGGACAAGGAAGATGGACACGATGCGCAGAAG 1715
```

Search completed: July 2, 2004, 12:46:11
Job time : 4046 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: July 2, 2004, 02:16:29 ; Search time 7788 Seconds
(without alignments)
7733.952 Million cell updates/sec

Title: US-10-051-902A-21
Perfect score: 2017
Sequence: 1 cttacatgtaagctcgtgcc.....aaaaaaaaaaaaaaaaaaaa 2017

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	796	39.5	804	14	CB621641	CB621641 OSIIEa07K
	2	764.4	37.9	768	14	CB620394	CB620394 OSIIEa05J
	3	763.2	37.8	824	14	CB621640	CB621640 OSIIEa07K
	4	746	37.0	818	14	CB625172	CB625172 OSIIEa14C
c	5	737.6	36.6	764	14	CB678733	CB678733 OSJNEf01D
	6	722.6	35.8	723	14	CB675155	CB675155 OSJNEe10K
	7	706.8	35.0	942	29	CG374317	CG374317 OGYBO78TV
	8	705	35.0	710	14	CB678732	CB678732 OSJNEf01D
c	9	687.4	34.1	697	14	CB620333	CB620333 OSIIEa05H
c	10	675.2	33.5	1006	29	CG283892	CG283892 OG1AC79TV
	11	656.4	32.5	671	14	CB620332	CB620332 OSIIEa05H
	12	654	32.4	700	9	AU057756	AU057756 AU057756
c	13	649	32.2	658	14	CB620395	CB620395 OSIIEa05J
c	14	648.6	32.2	929	29	CG260290	CG260290 OG3DF74TV
c	15	635	31.5	972	29	CG064683	CG064683 PUJCW07TD
c	16	625	31.0	641	14	CB648567	CB648567 OSJNEb12A
	17	621.8	30.8	853	29	CG189646	CG189646 PUJBO76TB
c	18	609.8	30.2	735	14	CB675156	CB675156 OSJNEe10K
	19	592.2	29.4	597	9	AU097469	AU097469 AU097469
	20	580	28.8	714	9	AU163952	AU163952 AU163952
	21	576.4	28.6	850	29	CG260280	CG260280 OG3DF74TH
c	22	576	28.6	884	28	CC378810	CC378810 PUHGF71TD
	23	575.8	28.5	689	9	AU095961	AU095961 AU095961
	24	575.8	28.5	852	29	CC731869	CC731869 OGUBY39TV
	25	556	27.6	868	29	CG041546	CG041546 PUJFB01TB
c	26	554.4	27.5	835	29	CG284028	CG284028 OGWGN33TH
c	27	552.6	27.4	926	29	CG374307	CG374307 OGYBO78TH
	28	538.8	26.7	823	29	CC645436	CC645436 OGUDD66TV
	29	523.4	25.9	920	28	CC443656	CC443656 PUHHP54TB
c	30	522.8	25.9	954	29	CC612757	CC612757 OGVBX53TH
c	31	520.4	25.8	786	29	CG222608	CG222608 OGXAE96TV
c	32	518.8	25.7	782	14	CB625173	CB625173 OSIIEa14C
c	33	515.8	25.6	802	29	CG309549	CG309549 OG0FE44TV
	34	515.2	25.5	699	12	BI810556	BI810556 K001C12 O
	35	512	25.4	736	28	BZ418680	BZ418680 if26a04.b
c	36	506	25.1	908	29	CG295842	CG295842 OG1DP44TV
	37	491.6	24.4	821	10	BF264405	BF264405 HV_CEa000
	38	489.2	24.3	736	14	CD882653	CD882653 F1.110K05
c	39	476.2	23.6	849	29	CG041551	CG041551 PUJFB01TD
	40	475.6	23.6	758	28	BZ635288	BZ635288 OGCAF60TC
c	41	467.8	23.2	882	28	CC348900	CC348900 OGUBM39TH
	42	464.4	23.0	902	29	CG283882	CG283882 OG1AC79TH
	43	463	23.0	800	14	CB623090	CB623090 OSIIEa10J
c	44	461.4	22.9	768	29	CG333772	CG333772 OG0CB43TH
	45	458.4	22.7	932	29	CG170800	CG170800 PUKBM54TB

ALIGNMENTS

RESULT 1
CB621641/c
LOCUS CB621641 804 bp mRNA linear EST 08-APR-2003
DEFINITION OSIIEa07K22.r OSIIEa Oryza sativa (indica cultivar-group) cDNA clone OSIIEa07K22 3', mRNA sequence.
ACCESSION CB621641
VERSION CB621641.1 GI:29616629
KEYWORDS EST.
SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 804)
AUTHORS Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E., Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
TITLE Large-scale identification of ESTs involved in the interaction between rice and Magnaporthe grisea
JOURNAL Unpublished (2003)
COMMENT Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ 85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu

```
            PCR PRimers
            FORWARD: gta aaa cga cgg cca gtg
            BACKWARD: gga aac agc tat gac cat g
            Plate: 07  row: K  column: 22
            Seq primer: gga aac agc tat gac cat g.
FEATURES             Location/Qualifiers
     source          1. .804
                     /organism="Oryza sativa (indica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="IR36"
                     /db_xref="taxon:39946"
                     /clone="OSIIEa07K22"
                     /tissue_type="Leaf"
                     /dev_stage="3 week"
                     /lab_host="DH10B"
                     /clone_lib="OSIIEa"
                     /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
                     XhoI; Lesion Mimic SPL 11"
ORIGIN

  Query Match             39.5%;  Score 796;  DB 14;  Length 804;
  Best Local Similarity   99.4%;  Pred. No. 2.6e-64;
  Matches  799;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1155 GGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCAC 1214
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        804 GGCGGGATGATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCAC 745

Qy       1215 CCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTC 1274
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        744 CCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTC 685

Qy       1275 GCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCG 1334
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        684 GCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCG 625

Qy       1335 CTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGC 1394
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        624 CTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGC 565

Qy       1395 GTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTC 1454
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        564 GTCATCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTTTTC 505

Qy       1455 CTCTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACC 1514
             || |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        504 CTTTACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTTTACACCTACCTCCCGGAGACC 445

Qy       1515 CGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAA 1574
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 CGCGGCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAA 385

Qy       1575 TCAGACGAGCCAGCCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCACTAACTGATCAAAC 1634
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 TCAGACGAGCCAGCCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCACTAACTGATCAAAC 325

Qy       1635 TAACCGCAAAATCACCAAATCCTAAGGGTTTTCTTGCAAAAACGTGTGCTGTACTGGCTA 1694
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        324 TAACCGCAAAATCACCAAATCCTAAGGGTTTTCTTGCGAAAACGTGTGCTGTACTGGCTA 265

Qy       1695 GCTAGCAAGTAGTAGCAGCAACGTGGGAAGATTCGCTGATCCGGCGTTGCTGGAGAGCGA 1754
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 GCTAGCAAGTAGTAGCAGCAACGTGGGAAGATTCGCTGATCCGGCGTTGCTGGAGAGCGA 205

Qy       1755 CGGCCGGCGACGACAAAGCTGAGCTCCAGCTCGAGACTTCTTAAAATCATCTTCAAGTAC 1814
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 CGGCCGGCGACGACAAAGCTGAGCTCCAGCTCGAGACTTCTTAAAATCATCTTCAAGTAC 145

Qy       1815 ATGGATTTTATTTTGCTCTTTGCTTTGTCCGTAAAAGTTGTACTATGCGATGAAGAATAC 1874
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 ATGGATTTTATTTTGCTCTTTGCTTTGTCCGTAAAAGTTGTACTATGCGATGAAGAATAC 85

Qy       1875 CAGTATGTAGCAAGGCTGAGGTTGTGTGTAGCTACTAGAAGTGTCAGTCACGTTGTTCTT 1934
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          84 CAGTATGTAGCAAGGCTGAGGTTGTGTGTAGCTACTAGAAGTGTCAGTCACGTTGTTCTT 25

Qy       1935 GTAAGAAATGTTTAACTGTTAATT 1958
             ||||||||||||||||||||||||
Db          24 GTAAGAAATGTTTAACTGTTAATT 1


RESULT 2
CB620394
LOCUS       CB620394                768 bp    mRNA    linear   EST 08-APR-2003
DEFINITION  OSIIEa05J08.f OSIIEa Oryza sativa (indica cultivar-group) cDNA
            clone OSIIEa05J08 5', mRNA sequence.
ACCESSION   CB620394
VERSION     CB620394.1  GI:29615382
KEYWORDS    EST.
SOURCE      Oryza sativa (indica cultivar-group)
  ORGANISM  Oryza sativa (indica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 768)
  AUTHORS   Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
            Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
  TITLE     Large-scale identification of ESTs involved in the interaction
            between rice and Magnaporthe grisea
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Rod Wing
            Arizona Genomics Institute
            University of Arizona
            Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
            85721-0088, USA
            Tel: 520 626 3967
            Fax: 520 621 9288
            Email: http://genome.arizona.edu
            PCR PRimers
            FORWARD: gta aaa cga cgg cca gtg
            BACKWARD: gga aac agc tat gac cat g
            Plate: 05  row: J  column: 08
            Seq primer: gta aaa cga cgg cca gtg.
FEATURES             Location/Qualifiers
     source          1. .768
                     /organism="Oryza sativa (indica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="IR36"
                     /db_xref="taxon:39946"
                     /clone="OSIIEa05J08"
                     /tissue_type="Leaf"
                     /dev_stage="3 week"
                     /lab_host="DH10B"
                     /clone_lib="OSIIEa"
                     /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
                     XhoI; Lesion Mimic SPL 11"
ORIGIN

  Query Match             37.9%;  Score 764.4;  DB 14;  Length 768;
  Best Local Similarity   99.6%;  Pred. No. 2.1e-61;
  Matches  765;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         41 CGCCACTACTGTACACGGCCCAGAGCGAGCCTCCTCCTCCTCTGCACCACCGGAGATGGC 100
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGCCACTACTGTCCACGGCCCAGAGCGAGCCTCCTCCTCCTCTGCACCACCGGAGATGGC 60

Qy        101 TTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCCTT 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTTCGCCTT 120

Qy        161 CGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTGAT 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGGGGTGAT 180

Qy        221 GAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAGGT 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        181 GAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGTGGAGGT 240
Qy        281 TCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTC 340
Db        241 TCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCGGACGTC 300
Qy        341 GGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGS 400
Db        301 GGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGCGGGGGC 360
Qy        401 GTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGG 460
Db        361 GTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGTGGCCGG 420
Qy        461 CATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGCCGGC 520
Db        421 CATCGGCGTGGGCTACGCGCTCATGATTGCGCCGGTGTACACCGCCGAGGTGTCGCCGGC 480
Qy        521 GTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCT 580
Db        481 GTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCCTGCT 540
Qy        581 CGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCAT 640
Db        541 CGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCATCAT 600
Qy        641 GCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGA 700
Db        601 GCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGA 660
Qy        701 GTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAA 760
Db        661 GTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAA 720
Qy        761 GACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAA 808
Db        721 GACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAA 768
```

RESULT 3
CB621640
LOCUS CB621640 824 bp mRNA linear EST 08-APR-2003
DEFINITION OSIIEa07K22.f OSIIEa Oryza sativa (indica cultivar-group) cDNA clone OSIIEa07K22 5', mRNA sequence.
ACCESSION CB621640
VERSION CB621640.1 GI:29616628
KEYWORDS EST.
SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 824)
AUTHORS Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E., Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
TITLE Large-scale identification of ESTs involved in the interaction between rice and Magnaporthe grisea
JOURNAL Unpublished (2003)
COMMENT Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ 85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR PRimers
FORWARD: gta aaa cga cgg cca gtg
BACKWARD: gga aac agc tat gac cat g
Plate: 07 row: K column: 22
Seq primer: gta aaa cga cgg cca gtg.
FEATURES Location/Qualifiers
source 1. .824
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="mRNA"
/cultivar="IR36"
/db_xref="taxon:39946"
/clone="OSIIEa07K22"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSIIEa"
/note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2: XhoI; Lesion Mimic SPL 11"
ORIGIN

Query Match 37.8%; Score 763.2; DB 14; Length 824;
Best Local Similarity 95.4%; Pred. No. 2.6e-61;
Matches 786; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

```
Qy        874 GAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCA 933
Db          1 GAAACGAGAAGCGCGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCA 60
Qy        934 TCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCG 993
Db         61 TCCTGCTGTCGGGCATCGGCATCCACTTCTTCCAGCAGGCGTCGGGCATCGACTCCGTCG 120
Qy        994 TCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGG 1053
Db        121 TGCTCTACAGCCCGCGCGTGTTCAAGAGCGCCGGCATCACCGACGACAACCACCTCCTCG 180
Qy       1054 GCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCA 1113
Db        181 GCACCACCTGCGCCGTCGGTGTCACCAAGACGCTCTTCATCCTCGTGGCGACCTTCTTCC 240
Qy       1114 TCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCC 1173
Db        241 TCGACCGCGTCGGGCGGCGGCCGCTGCTGCTGAGCAGCACGGGCGGGATGATCCTCTCCC 300
Qy       1174 TCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTT 1233
Db        301 TCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTT 360
Qy       1234 GGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCC 1293
Db        361 GGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCC 420
Qy       1294 TTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGG 1353
Db        421 TTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGG 480
Qy       1354 GCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCC 1413
Db        481 GCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCC 540
Qy       1414 TGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCG 1473
Db        541 TGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCG 600
Qy       1474 CGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGG 1533
Db        601 CGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGG 660
Qy       1534 AGATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAGCCAAGG 1593
Db        661 AGATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAGCCAAGG 720
Qy       1594 AGAAGAAGAAGGTGGAAATGGCCGCCACTAACTGATCAAACTAACCGCAAAATCACCAAA 1653
Db        721 AGAAGAAGAAGGTGGAAATGGCCGCCACTAACTGATCAAACTAACCGCAAAATCACCAAA 780
Qy       1654 TCCTAAGGGTTTTCTTGCAAAAACGTGTGCTGTACTGGCTAGCT 1697
Db        781 TCCTAAGGGTTTTCTTGCGAAAACGTGTGCTGTACTGGCTAGCT 824
```

```
RESULT 4
CB625172
LOCUS       CB625172                818 bp    mRNA    linear   EST 08-APR-2003
DEFINITION  OSIIEa14C12.f OSIIEa Oryza sativa (indica cultivar-group) cDNA
            clone OSIIEa14C12 5', mRNA sequence.
ACCESSION   CB625172
VERSION     CB625172.1  GI:29620160
KEYWORDS    EST.
SOURCE      Oryza sativa (indica cultivar-group)
  ORGANISM  Oryza sativa (indica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 818)
  AUTHORS   Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
            Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
  TITLE     Large-scale identification of ESTs involved in the interaction
            between rice and Magnaporthe grisea
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Rod Wing
            Arizona Genomics Institute
            University of Arizona
            Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
            85721-0088, USA
            Tel: 520 626 3967
            Fax: 520 621 9288
            Email: http://genome.arizona.edu
            PCR PRimers
            FORWARD: gta aaa cga cgg cca gtg
            BACKWARD: gga aac agc tat gac cat g
            Plate: 14  row: C  column: 12
            Seq primer: gta aaa cga cgg cca gtg.
FEATURES             Location/Qualifiers
     source          1. .818
                     /organism="Oryza sativa (indica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="IR36"
                     /db_xref="taxon:39946"
                     /clone="OSIIEa14C12"
                     /tissue_type="Leaf"
                     /dev_stage="3 week"
                     /lab_host="DH10B"
                     /clone_lib="OSIIEa"
                     /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
                     XhoI; Lesion Mimic SPL 11"
ORIGIN

  Query Match             37.0%;  Score 746;  DB 14;  Length 818;
  Best Local Similarity   94.5%;  Pred. No. 9.7e-60;
  Matches  773;  Conservative    0;  Mismatches   45;  Indels    0;  Gaps    0;

Qy        457 CCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGC 516
              ||||||||||||||||| || ||||||||||||||||||||||||||||||||||||||||
Db          1 CCGGCATCGGCGTGGGCCACTCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTCGC 60

Qy        517 CGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCC 576
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCATCC 120

Qy        577 TGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCA 636
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCGCA 180

Qy        637 TCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGC 696
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGC 240

Qy        697 CGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGG 756
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGG 300

Qy        757 AGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCG 816
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCG 360

Qy        817 CCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAA 876
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCGGCATCCCCGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGGAA 420

Qy        877 ACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCC 936
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCC 480

Qy        937 TGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCT 996
              ||||||| || |||||||||||||||||||||| |||| |||||| ||||||||||
Db        481 TGCTGTCGGGCATCGGCATCCACTTCTTCCAGCAGGCGTCGGGCATCGACTCCGTCGTGC 540

Qy        997 TCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGCA 1056
              |||||||||| | |||||||||||||| |||| | || ||||||| ||| || | ||||
Db        541 TCTACAGCCCGCGCGTGTTCAAGAGCGCCGGCATCACCGACGACAACCACCTCCTCGGCA 600

Qy       1057 CCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCG 1116
              |||| ||  |  |||||||||||||| ||| |||||| |  ||||||| |||||| |||
Db        601 CCACCTGCGCCGTCGGTGTCACCAAGACGCTCTTCATCCTCGTGGCGACCTTCTTCCTCG 660

Qy       1117 ACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCA 1176
              || |||||||||||||||||| ||||| |||||||||||||| ||||||||||||
Db        661 ACCGCGTCGGGCGGCGGCCGCTGCTGCTGAGCAGCACGGGCGGGATGATCCTCTCCCTCA 720

Qy       1177 TCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGG 1236
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGG 780

Qy       1237 CCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTC 1274
              |||||||||||||||||||||| ||   |||||||
Db        781 CCATCGGCCTAAGCATCGCCTCCCACCCTCGCTACGTC 818


RESULT 5
CB678733/c
LOCUS       CB678733                764 bp    mRNA    linear   EST 09-APR-2003
DEFINITION  OSJNEf01D24.r OSJNEf Oryza sativa (japonica cultivar-group) cDNA
            clone OSJNEf01D24 3', mRNA sequence.
ACCESSION   CB678733
VERSION     CB678733.1  GI:29682458
KEYWORDS    EST.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 764)
  AUTHORS   Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
            Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
  TITLE     Large-scale identification of ESTs involved in the interaction
            between rice and Magnaporthe grisea
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Rod Wing
            Arizona Genomics Institute
            University of Arizona
            Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
            85721-0088, USA
            Tel: 520 626 3967
            Fax: 520 621 9288
            Email: http://genome.arizona.edu
            PCR PRimers
            FORWARD: gta aaa cga cgg cca gtg
            BACKWARD: gga aac agc tat gac cat g
            Plate: 01  row: D  column: 24
            Seq primer: gga aac agc tat gac cat g.
FEATURES             Location/Qualifiers
     source          1. .764
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="Nipponbare"
```

```
                     /db_xref="taxon:39947"
                     /clone="OSJNEf01D24"
                     /tissue_type="Leaf"
                     /dev_stage="3 week"
                     /lab_host="DH10B"
                     /clone_lib="OSJNEf"
                     /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
                     XhoI; Uninfected Control"
ORIGIN

  Query Match             36.6%;  Score 737.6;  DB 14;  Length 764;
  Best Local Similarity   98.6%;  Pred. No. 5.9e-59;
  Matches  755;  Conservative    0;  Mismatches    9;  Indels    2;  Gaps    1;

Qy       1219 ACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCT 1278
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        764 ACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCT 705

Qy       1279 TCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCC 1338
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        704 TCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCC 645

Qy       1339 AGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCA 1398
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        644 AGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCA 585

Qy       1399 TCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCT 1458
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        584 TCTCCATGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCT 525

Qy       1459 ACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCG 1518
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        524 ACTCCGGCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCG 465

Qy       1519 GCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAG 1578
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        464 GCCGGACGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAG 405

Qy       1579 ACGAGCCAGCCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCACTAACTGATCAAACTAAC 1638
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        404 ACGAGCCAGCCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCACTAACTGATCAAACTAAC 345

Qy       1639 CGCAAAATCACCAAATCCTAAGGGTTTTCTTGCAAAAACGTGTGCTGTACTGGCTAGCTA 1698
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        344 CGCAAAATCACCAAATCCTAAGGGTTTTCTTGCAAAAACGTGTGCTGTACTGGCTAGCTA 285

Qy       1699 GCAAGTAGTAGCAGCAACGTGGGAAGATTCGCTGATCCGGCGTTGCTGGAGAGCGACGGC 1758
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 GCAAGTAGTAGCAGCAACGTGGGAAGATTCGCTGATCCGGCGTTGCTGGAGAGCGACGGC 225

Qy       1759 CGGCGACGACAAAGCTGAGCTCCAGCTCGAGACTTCTTAAAATCATCTTCAAGTACATGG 1818
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 CGGCGACGACAAAGCTGAGCTCCAGCTCGAGACTTCTTAAAATCATCTTCAAGTACATGG 165

Qy       1819 ATTTTATTTTGCTCTTTGCTTTGTCCGTAAAAGTTGTACTATGCGATGAAGAATACCAGT 1878
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        164 ATTTTATTTTGCTCTTTGCTTTGTCCGTAAAAGTTGTACTATGCGATGAAGAATACCAGT 105

Qy       1879 ATGTAGCAAGGCTGAGGTTGTGTGTAGCTACTAGAAGTGTCAGTCACGTTGTTCTTGTAA 1938
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        104 ATGTAGCAAGGCTGAGGTTGTGTGTAGCTACTAGAAGTGTCAGTCACGTTGTTCTTGTAA 45

Qy       1939 GAAATGTTTAACTGTTAATTAAGCAGTATTGTTGCAGTAATCAAAA 1984
             ||||||     |    ||||||||||||||||||||||||   |||
Db         44 GAAATG--AGAACTGTAATTAAGCAGTATTGTTGCAGTAAATTAAA 1

RESULT 6
CB675155
LOCUS       CB675155                723 bp    mRNA    linear   EST 09-APR-2003
DEFINITION  OSJNEe10K23.f OSJNEe Oryza sativa (japonica cultivar-group) cDNA
            clone OSJNEe10K23 5', mRNA sequence.
ACCESSION   CB675155
VERSION     CB675155.1  GI:29678880
KEYWORDS    EST.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 723)
  AUTHORS   Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
            Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
  TITLE     Large-scale identification of ESTs involved in the interaction
            between rice and Magnaporthe grisea
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Rod Wing
            Arizona Genomics Institute
            University of Arizona
            Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
            85721-0088, USA
            Tel: 520 626 3967
            Fax: 520 621 9288
            Email: http://genome.arizona.edu
            PCR PRimers
            FORWARD: gta aaa cga cgg cca gtg
            BACKWARD: gga aac agc tat gac cat g
            Plate: 10  row: K  column: 23
            Seq primer: gta aaa cga cgg cca gtg.
FEATURES             Location/Qualifiers
     source          1. .723
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /clone="OSJNEe10K23"
                     /tissue_type="Leaf"
                     /dev_stage="3 week"
                     /lab_host="DH10B"
                     /clone_lib="OSJNEe"
                     /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
                     XhoI; 24 hrs after innoculation with Rice Blast (70-15)"
ORIGIN

  Query Match             35.8%;  Score 722.6;  DB 14;  Length 723;
  Best Local Similarity   99.9%;  Pred. No. 1.4e-57;
  Matches  722;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         35 CTCGACCGCCACTACTGTACACGGCCCAGAGCGAGCCTCCTCCTCCTCTGCACCACCGGA 94
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTCGACCGCCACTACTGTACACGGCCCAGAGCGAGCCTCCTCCTCCTCTGCACCACCGGA 60

Qy         95 GATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTT 154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGGTT 120

Qy        155 CGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGG 214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATCGG 180

Qy        215 GGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGT 274
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAGGT 240

Qy        275 GGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCG 334
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGGCG 300

Qy        335 GACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGC 394
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTCGC 360

Qy        395 GGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGT 454
              |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGGGGCGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTCGT 420
```

```
Qy        455 GGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTC 514
Db        421 GGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTGTC 480

Qy        515 GCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCAT 574
Db        481 GCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGCAT 540

Qy        575 CCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCG 634
Db        541 CCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGGCG 600

Qy        635 CATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCAT 694
Db        601 CATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCAT 660

Qy        695 GCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCT 754
Db        661 GCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCT 720

Qy        755 GGA 757
Db        721 GGA 723
```

RESULT 7
CG374317
LOCUS CG374317 942 bp DNA linear GSS 26-AUG-2003
DEFINITION OGYBO78TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0642M11, genomic survey sequence.
ACCESSION CG374317
VERSION CG374317.1 GI:34291584
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 942)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
TITLE Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Other_GSSs: OGYBO78TH
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
FEATURES Location/Qualifiers
source 1. .942
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMa0642M11"
/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"
ORIGIN

Query Match 35.0%; Score 706.8; DB 29; Length 942;
Best Local Similarity 84.4%; Pred. No. 3.6e-56;
Matches 795; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

```
Qy        434 GCTCATGTTCGGCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCC 493
Db          1 GCTCATGTTCGGGCGCTTCGTGGCCGGGATCGGCGTGGGCTACGCGCTCATGATCGCGCC 60

Qy        494 GGTGTACACCGCCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGA 553
Db         61 CGTGTACACGGCCGAGGTGTCCCCCGCGTCGGCCCGCGGCTTCCTGACGTCGTTCCCGGA 120

Qy        554 GGTGTTCATCAACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTT 613
Db        121 GGTGTTCATCAACTTCGGCATCCTCCTAGGCTACGTGTCCAACTACGCCTTCTCCCGCCT 180

Qy        614 GCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCT 673
Db        181 CCCGCTCCGCCTGGGCTGGCGCGTGATGCTCGGCATCGGCGCGGCCCCGTCCGTGGTGCT 240

Qy        674 CGCGCTCATGGTGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCT 733
Db        241 CGCCCTCATGGTGCTCGGCATGCCCGAGTCGCCCCGGTGGCTCGTCATGAAGGGCCGCCT 300

Qy        734 CGCGGACGCCAAGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCG 793
Db        301 GGCGGACGCCAAGGTGGTGCTGGGGAAGACCTCCGACACGCCGGAGGAGGCCGCGGAGAG 360

Qy        794 CCTGGCCGACATCAAGGCCGCCGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGAC 853
Db        361 GCTCGCTGACATCAAGGCCGCCGCCGGCATCCCGCAGGAGCTCGACGGCGACGTGGTGGC 420

Qy        854 CGTCCCCAAGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCC 913
Db        421 CGTGCCAAAGAAGCGGGACAGCGAGGAGGCGCGGGTGTGGAAGGAGCTCATCCTGTCCCC 480

Qy        914 GACCCCGGCCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGC 973
Db        481 GACCCCGGCCATGCGTCGCATCCTCCTCTCGGGCATCGGCATCCACTTCTTCCAGCAGGC 540

Qy        974 GTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAAC 1033
Db        541 GTCCGGCATCGACGCCGTGGTGCTGTACAGCCCGCGCGTGTTCAAGGCCGCCGGCATCAC 600

Qy       1034 GAACGACAAACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCAT 1093
Db        601 CAGCGACAACCAGCTCCTGGGAACGACCTGCGCGGTGGGCGTCACCAAGACGCTCTTCAT 660

Qy       1094 CTTGTTGGCGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCAC 1153
Db        661 CCTGGTGGCCACGTTCCTGCTCGATAGAGTCGGCCGGCGGCCGCTCCTCCTGAGCAGCGT 720

Qy       1154 GGGCGGGATAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCA 1213
Db        721 GGGCGGCATGATCTTCTCGCTGGTCGGGCTCGCCGCCGGGCTCACCGTCATCGGGCACTA 780

Qy       1214 CCCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGT 1273
Db        781 CCAGGACGAAAAGATCCCATGGGCCATCGGCGTGGCCATCGCGTCTACCATGGCCTACGT 840

Qy       1274 CGCCTTCTTCTCCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCC 1333
Db        841 GGCCTTCTTCTCCATCGGCCTCGGCCCCATCACGTGGGTGTACAGCTCGGAGGTATTCCC 900

Qy       1334 GCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCGTCGCCGC 1375
Db        901 GCTGCATGTGCGCGCGATGGGCTGCGCGCTGGGCGTGGCCTC 942
```

RESULT 8
CB678732
LOCUS CB678732 710 bp mRNA linear EST 09-APR-2003
DEFINITION OSJNEf01D24.f OSJNEf Oryza sativa (japonica cultivar-group) cDNA clone OSJNEf01D24 5', mRNA sequence.
ACCESSION CB678732
VERSION CB678732.1 GI:29682457
KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

```
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 710)
  AUTHORS   Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
            Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
  TITLE     Large-scale identification of ESTs involved in the interaction
            between rice and Magnaporthe grisea
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Rod Wing
            Arizona Genomics Institute
            University of Arizona
            Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
            85721-0088, USA
            Tel: 520 626 3967
            Fax: 520 621 9288
            Email: http://genome.arizona.edu
            PCR PRimers
            FORWARD: gta aaa cga cgg cca gtg
            BACKWARD: gga aac agc tat gac cat g
            Plate: 01  row: D  column: 24
            Seq primer: gta aaa cga cgg cca gtg.
FEATURES             Location/Qualifiers
     source          1. .710
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /clone="OSJNEf01D24"
                     /tissue_type="Leaf"
                     /dev_stage="3 week"
                     /lab_host="DH10B"
                     /clone_lib="OSJNEf"
                     /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
                     XhoI; Uninfected Control"
ORIGIN

  Query Match           35.0%;  Score 705;  DB 14;  Length 710;
  Best Local Similarity  99.7%;  Pred. No. 5.9e-56;
  Matches  705;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         33 CACTCGACCGCCACTACTGTACACGGCCCAGAGCGAGCCTCCTCCTCCTCTGCACCACCG  92
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 CTCTCGACCGCCACTACTGTACACGGCCCAGAGCGAGCCTCCTCCTCCTCTGCACCACCG  63

Qy         93 GAGATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGG 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 GAGATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGG 123

Qy        153 TTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATC 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 TTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATC 183

Qy        213 GGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAG 272
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 GGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAG 243

Qy        273 GTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGG 332
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 GTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGG 303

Qy        333 CGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTC 392
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 CGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTC 363

Qy        393 GCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTC 452
              |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 GCGGGGGCGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTC 423

Qy        453 GTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTG 512
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 GTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTG 483

Qy        513 TCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGC 572
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 TCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGC 543

Qy        573 ATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGG 632
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 ATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGG 603

Qy        633 CGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGC 692
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 CGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGC 663

Qy        693 ATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGA 739
              |||||||||||||||||||||||||||||||||||||||||||||
Db        664 ATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGA 710


RESULT 9
CB620333/c
LOCUS       CB620333                697 bp    mRNA    linear   EST 08-APR-2003
DEFINITION  OSIIEa05H07.r OSIIEa Oryza sativa (indica cultivar-group) cDNA
            clone OSIIEa05H07 3', mRNA sequence.
ACCESSION   CB620333
VERSION     CB620333.1  GI:29615321
KEYWORDS    EST.
SOURCE      Oryza sativa (indica cultivar-group)
  ORGANISM  Oryza sativa (indica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 697)
  AUTHORS   Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
            Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
  TITLE     Large-scale identification of ESTs involved in the interaction
            between rice and Magnaporthe grisea
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Rod Wing
            Arizona Genomics Institute
            University of Arizona
            Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
            85721-0088, USA
            Tel: 520 626 3967
            Fax: 520 621 9288
            Email: http://genome.arizona.edu
            PCR PRimers
            FORWARD: gta aaa cga cgg cca gtg
            BACKWARD: gga aac agc tat gac cat g
            Plate: 05  row: H  column: 07
            Seq primer: gga aac agc tat gac cat g.
FEATURES             Location/Qualifiers
     source          1. .697
                     /organism="Oryza sativa (indica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="IR36"
                     /db_xref="taxon:39946"
                     /clone="OSIIEa05H07"
                     /tissue_type="Leaf"
                     /dev_stage="3 week"
                     /lab_host="DH10B"
                     /clone_lib="OSIIEa"
                     /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
                     XhoI; Lesion Mimic SPL 11"
ORIGIN

  Query Match           34.1%;  Score 687.4;  DB 14;  Length 697;
  Best Local Similarity  99.1%;  Pred. No. 2.4e-54;
  Matches  691;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1234 GGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCC 1293
              |||||||||| ||||||||||||||||||||| |||||||||||||||||||||||||||
Db        697 GGGCCATCGGCGTAAGCATCGCCTCCACCCTCGCATACGTCGCCTTCTTCTCCATCGGCC 638

Qy       1294 TTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGG 1353
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        637 TTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGG 578
```

```
Qy       1354 GCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCC 1413
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        577 GCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCC 518

Qy       1414 TGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCG 1473
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        517 TGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCG 458

Qy       1474 CGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGG 1533
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        457 CGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGG 398

Qy       1534 AGATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAGCCAAGG 1593
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        397 AGATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAGCCAAGG 338

Qy       1594 AGAAGAAGAAGGTGGAAATGGCCGCCACTAACTGATCAAACTAACCGCAAAATCACCAAA 1653
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        337 AGAAGAAGAAGGTGGAAATGGCCGCCACTAACTGATCAAACTAACCGCAAAATCACCAAA 278

Qy       1654 TCCTAAGGGTTTTCTTGCAAAAACGTGTGCTGTACTGGCTAGCTAGCAAGTAGTAGCAGC 1713
             |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        277 TCCTAAGGGTTTTCTTGCGAAAACGTGTGCTGTACTGGCTAGCTAGCAAGTAGTAGCAGC 218

Qy       1714 AACGTGGGAAGATTCGCTGATCCGGCGTTGCTGGAGAGCGACGGCCGGCGACGACAAAGC 1773
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 AACGTGGGAAGATTCGCTGATCCGGCGTTGCTGGAGAGCGACGGCCGGCGACGACAAAGC 158

Qy       1774 TGAGCTCCAGCTCGAGACTTCTTAAAATCATCTTCAAGTACATGGATTTTATTTTGCTCT 1833
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         157 TGAGCTCCAGCTCGAGACTTCTTAAAATCATCTTCAAGTACATGGATTTTATTTTGCTCT 98

Qy       1834 TTGCTTTGTCCGTAAAAGTTGTACTATGCGATGAAGAATACCAGTATGTAGCAAGGCTGA 1893
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          97 TTGCTTTGTCCGTAAAAGTTGTACTATGCGATGAAGAATACCAGTATGTAGCAAGGCTGA 38

Qy       1894 GGTTGTGTGTAGCTACTAGAAGTGTCAGTCACGTTGT 1930
             ||||||||| || |||||||||||||| |||||||||
Db          37 GGTTGTGTGAAGGTACTAGAAGTGTCATTCACGTTGT 1
```

RESULT 10
CG283892/c
LOCUS CG283892 1006 bp DNA linear GSS 25-AUG-2003
DEFINITION OG1AC79TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0715N13, genomic survey sequence.
ACCESSION CG283892
VERSION CG283892.1 GI:34198106
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 1006)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
TITLE Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Other_GSSs: OG1AC79TH
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
FEATURES Location/Qualifiers
source 1. .1006
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMa0715N13"
/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"
ORIGIN

Query Match 33.5%; Score 675.2; DB 29; Length 1006;
Best Local Similarity 80.2%; Pred. No. 2.7e-53;
Matches 804; Conservative 1; Mismatches 194; Indels 3; Gaps 1;

```
Qy        205 ACGATATCGGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTG 264
              | || |||||||||||||||||||| |||||||||||  |||||| | || |||| ||
Db       1002 AAGACATCGGGGTGATGAGCGGGGCGGCGCTGTACATCAGAAAGGACCTGAAGATCACTG 943

Qy        265 ACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCG 324
              ||| | || ||||| | || |||||||| || ||| ||||||||||||||||||||||||
Db        942 ACGTGCAGCTGGAGATCCTGATGGGCATCCTCAACTTCTACTCGCTCATCGGCTCCTTCG 883

Qy        325 CGGCGGGGCGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCA 384
              |||||||||||||||| |||||||||| || |||| ||| |||| |||| ||||
Db        882 CGGCGGGGCGGACGTCTGACTGGATCGGTCGCCGGTTCACTGTCGTCGTCGCGGCCGCGT 823

Qy        385 TATTCTTCGCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCG 444
              |||||||||| || |:  | |||||||  ||||||  |  |||||||| ||| ||| | |
Db        822 TATTCTTCGCCGGCGCCCTGCTCATGGGCCTCGCCGGCGGCTACGCCACGCTGATGCTGG 763

Qy        445 GCCGCTTCGTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCG 504
              |||||||||||||||| |||||||||||||| ||| ||||||||||||||||||||||
Db        762 GCCGCTTCGTGGCCGGCGTCGGCGTGGGCTACGGGCTGATGATCGCGCCGGTGTACACCG 703

Qy        505 CCGAGGTGTCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCA 564
              |||||||||| || |||||||| || |||||||| || || ||| | || ||||||||||
Db        702 CCGAGGTGTCCCCCGCGTCGGCCCGCGGCTTCCTCACCTCCTTCTCCGAAGTGTTCATCA 643

Qy        565 ACTTCGGCATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACC 624
              || |||| |||| |||||||| |||||| || ||| | ||| |||||||    ||
Db        642 ACGTCGGGATCCTCCTCGGCTACGTCTCCAACTACGCCTTCGCACGCCTGCCGCTCCGCC 583

Qy        625 TCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGG 684
              |||| |||||| |||||||||||||| |||||||||||| ||||| ||||||||||
Db        582 TCGGCTGGCGCGTCATGCTCGGCATCGGAGCGGCGCCGTCCGTCCTGCTGGCGCTCATGG 523

Qy        685 TGCTCGGCATGCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCA 744
              || |||||||||| ||||| || || ||||| |||||||||| ||||||||||||||||
Db        522 TGTTCGGCATGCCCGAGTCCCCCCGCTGGCTCGTCATGAAGGGCCGCCTCGCGGACGCCA 463

Qy        745 AGGTGGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACA 804
              ||| |||||||| || | | |||| ||| |||||||||| ||||| ||||||||||||||
Db        462 GGGTCGTGCTGGACAAAATCGCCGAGACGCCGGAGGAGGCTGCGGAACGCCTGGCCGACA 403

Qy        805 TCAAGGCCGCCGCCGGCATCCC---TGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCA 861
              |||||||||||||||| |||||    || |    ||| |||||||||||   ||||  |||
Db        402 TCAAGGCCGCCGCCGGAATCCCCCGCGAAGGAGCCGATGGCGACGTGGTCGCCGTGGCCA 343

Qy        862 AGAGAGGGAGCGGAAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGG 921
              ||||   |||||   | |||| |||||| |||||||||| ||||||||||||||  |||
Db        342 AGAGGAAGAGCGCCGAGGAGAGGCGGGTCTGGAAGGAGCTGATCCTGTCCCCGACGCCGA 283

Qy        922 CCATGCGGCGCATCCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCA 981
              || |||||||||||||||  |  | | ||  |||| |||||||||||  || |  |||
Db        282 CCGTGCGGCGCATCCTGCTCGCGACGCTTGGTCTCCAGTTCTTCCAGCAGGCCTCAGGCG 223

Qy        982 TTCACTCCGTCGTCTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACA 1041
              || |||||||||| | |||||||| | ||||||| |||| | ||  | ||  | |||
Db        222 TTGACTCCGTCGTGATGTACAGCCCCCGCGTGTTCCAGAGCGCGGGCATCACCGATGACG 163

Qy       1042 AACACTTCTTGGGCACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGG 1101
              | ||  | | |||||| ||  |  | || |||||||||| | ||  || | |||||||
Db        162 ACCAGCTTCTTGCCACCACCTGCGCCGTGGGCGTCACCAAGACGGTTATCGTGTTGTTGG 103
```

```
Qy       1102 CGACTTTCTTCATCGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGA 1161
Db        102 CCGCGGTCCTCCTGGACCGCGTCGGCAGGCGCCCGCTGATGCTGACCAGCTCCGGCGGCA 43

Qy       1162 TAATCCTCTCCCTCATCGGCCTCGGCGCCGGGCTCACCGTCG 1203
Db         42 TGGCTGTCTCGCTCGTCGGCCTCGCCACAGGGCTCACCGTCG 1
```

RESULT 11
CB620332
LOCUS CB620332 671 bp mRNA linear EST 08-APR-2003
DEFINITION OSIIEa05H07.f OSIIEa Oryza sativa (indica cultivar-group) cDNA clone OSIIEa05H07 5', mRNA sequence.
ACCESSION CB620332
VERSION CB620332.1 GI:29615320
KEYWORDS EST.
SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 671)
AUTHORS Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E., Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
TITLE Large-scale identification of ESTs involved in the interaction between rice and Magnaporthe grisea
JOURNAL Unpublished (2003)
COMMENT Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ 85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR PRimers
FORWARD: gta aaa cga cgg cca gtg
BACKWARD: gga aac agc tat gac cat g
Plate: 05 row: H column: 07
Seq primer: gta aaa cga cgg cca gtg.

FEATURES Location/Qualifiers
source 1. .671
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="mRNA"
/cultivar="IR36"
/db_xref="taxon:39946"
/clone="OSIIEa05H07"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSIIEa"
/note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2: XhoI; Lesion Mimic SPL 11"

ORIGIN

Query Match 32.5%; Score 656.4; DB 14; Length 671;
Best Local Similarity 98.8%; Pred. No. 1.7e-51;
Matches 660; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

```
Qy         33 CACTCGACCGCCACTACTGTACACGGCCCAGAGCGAGCCTCCTCCTCCTCTGCACCACCG 92
Db          4 CTCTCGACCCCCTTAAGTGTACACGGCCCAGAGCGAGCCTCCTCCTCCTCTGCACCACCG 63

Qy         93 GAGATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCGAAGAAGAAGGGCAACGTCCGG 152
Db         64 GAGATGGCTTCCGCCGCGCTGCCGGAGGCCGTCGCGCCTAAGAAGAAGGGCAACGTCCGG 123

Qy        153 TTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATC 212
Db        124 TTCGCCTTCGCCTGCGCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTACGATATC 183

Qy        213 GGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAG 272
Db        184 GGGGTGATGAGCGGGGCGTCGCTGTACATCAAGAAGGACTTCAACATCAGTGACGGGAAG 243

Qy        273 GTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGG 332
Db        244 GTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTCGCGGCGGGG 303

Qy        333 CGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTC 392
Db        304 CGGACGTCGGACTGGATCGGCCGGCGGTACACCATCGTGTTCGCCGCCGTCATATTCTTC 363

Qy        393 GCGGGGGSGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTC 452
Db        364 GCGGGGGCGTTCCTCATGGGGTTCGCCGTCAACTACGCCATGCTCATGTTCGGCCGCTTC 423

Qy        453 GTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTG 512
Db        424 GTGGCCGGCATCGGCGTGGGCTACGCGCTCATGATCGCGCCGGTGTACACCGCCGAGGTG 483

Qy        513 TCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGC 572
Db        484 TCGCCGGCGTCGGCGCGTGGCTTCCTGACGTCGTTCCCGGAGGTGTTCATCAACTTCGGC 543

Qy        573 ATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGG 632
Db        544 ATCCTGCTCGGGTACGTCTCGAACTATGCTTTCTCCCGCTTGCCGCTGAACCTCGGGTGG 603

Qy        633 CGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGC 692
Db        604 CGCATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGC 663

Qy        693 ATGCCGGA 700
Db        664 ATGCCGGA 671
```

RESULT 12
AU057756
LOCUS AU057756 700 bp mRNA linear EST 01-APR-2002
DEFINITION AU057756 Oryza sativa mature leaf Nipponbare Oryza sativa (japonica cultivar-group) cDNA clone S21759_1A, mRNA sequence.
ACCESSION AU057756
VERSION AU057756.1 GI:4716640
KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 700)
AUTHORS Yamamoto,K. and Sasaki,T.
TITLE Rice cDNA from mature leaf
JOURNAL Unpublished (1999)
COMMENT Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
Tel: 81-298-38-7441
Fax: 81-298-38-7468
Email: tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/
PROJECT ='RGP'.

FEATURES Location/Qualifiers
source 1. .700
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="S21759_1A"
/tissue_type="mature leaf"
/clone_lib="Oryza sativa mature leaf Nipponbare"

ORIGIN

Query Match 32.4%; Score 654; DB 9; Length 700;

```
  Best Local Similarity   99.4%;  Pred. No. 2.8e-51;
  Matches  698;  Conservative    0;  Mismatches    0;  Indels    4;  Gaps    4;

Qy       1245 CTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATC 1304
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGGCCCCATC 60

Qy       1305 ACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTC 1364
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTC 120

Qy       1365 GGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCC 1424
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTCGCTGTCC 180

Qy       1425 AAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGG 1484
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCTCGCCTGG 240

Qy       1485 GTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAG 1544
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGATGAGCAAG 300

Qy       1545 CTGTTCGGCGACACGGCCGCCGCC-TCGGAATCAGACGAGCCAGCCAAGGAGAAGAAGAA 1603
             |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        301 CTGTTCGGCGACACGGCCGCCGCCTTCGGAATCAGACGAGCCAGCCAAGGAGAAGAAGAA 360

Qy       1604 GGTGGAAATGGCCGCCACTAACTGATCAAACTAACCGCAAAATCACCAAATCCTAAGGGT 1663
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGTGGAAATGGCCGCCACTAACTGATCAAACTAACCGCAAAATCACCAAATCCTAAGGGT 420

Qy       1664 TTTCTTGCAAAAACGTGTGCTGTACTGGCTAGCTAGCAAGTAGTAGCAGCAACGTGGGAA 1723
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TTTCTTGCAAAAACGTGTGCTGTACTGGCTAGCTAGCAAGTAGTAGCAGCAACGTGGGAA 480

Qy       1724 GATTCGCTGATCCGGCGTTGCTGGAGAGCGACGGCCGGCGACGACAAAGCTGAGCTCCAG 1783
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GATTCGCTGATCCGGCGTTGCTGGAGAGCGACGGCCGGCGACGACAAAGCTGAGCTCCAG 540

Qy       1784 CTCGAGACTTCTTAAAATCATCTTCAAGTACATGGATTTT-ATTTTGCTCTTTGCTTTGT 1842
             ||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db        541 CTCGAGACTTCTTAAAATCATCTTCAAGTACATGGATTTTAATTTTGCTCTTTGCTTTGT 600

Qy       1843 CCGTAAAAGTTGTACTATGCGATGAAGAATACCAGTATGTAGCAAGGCTGAGGTTGTGTG 1902
             ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        601 CCGTAAAAGTTGTACTATGCGATGAAG-ATACCAGTATGTAGCAAGGCTGAGGTTGTGTG 659

Qy       1903 TAGCTACTAGAAGTGTCAGTCACGTTGTTCTTGTAAGAAATG 1944
             ||| |||||||||||||||||||||||||||||||||||||||
Db        660 TAG-TACTAGAAGTGTCAGTCACGTTGTTCTTGTAAGAAATG 700


RESULT 13
CB620395/c
LOCUS       CB620395                658 bp    mRNA    linear   EST 08-APR-2003
DEFINITION  OSIIEa05J08.r OSIIEa Oryza sativa (indica cultivar-group) cDNA
            clone OSIIEa05J08 3', mRNA sequence.
ACCESSION   CB620395
VERSION     CB620395.1  GI:29615383
KEYWORDS    EST.
SOURCE      Oryza sativa (indica cultivar-group)
  ORGANISM  Oryza sativa (indica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 658)
  AUTHORS   Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
            Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
  TITLE     Large-scale identification of ESTs involved in the interaction
            between rice and Magnaporthe grisea
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Rod Wing
            Arizona Genomics Institute
            University of Arizona
            Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
            85721-0088, USA
            Tel: 520 626 3967
            Fax: 520 621 9288
            Email: http://genome.arizona.edu
            PCR PRimers
            FORWARD: gta aaa cga cgg cca gtg
            BACKWARD: gga aac agc tat gac cat g
            Plate: 05  row: J  column: 08
            Seq primer: gga aac agc tat gac cat g.
FEATURES             Location/Qualifiers
     source          1. .658
                     /organism="Oryza sativa (indica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="IR36"
                     /db_xref="taxon:39946"
                     /clone="OSIIEa05J08"
                     /tissue_type="Leaf"
                     /dev_stage="3 week"
                     /lab_host="DH10B"
                     /clone_lib="OSIIEa"
                     /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
                     XhoI; Lesion Mimic SPL 11"
ORIGIN

  Query Match             32.2%;  Score 649;  DB 14;  Length 658;
  Best Local Similarity   99.1%;  Pred. No. 8.3e-51;
  Matches  652;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1285 CCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGC 1344
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         658 CCATCGGCCTTGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGC 599

Qy       1345 GCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCA 1404
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         598 GCGCGCTGGGCTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCA 539

Qy       1405 TGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCG 1464
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         538 TGACCTTCCTGTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCG 479

Qy       1465 GCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGA 1524
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         478 GCATCGCCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGA 419

Qy       1525 CGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGC 1584
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         418 CGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGC 359

Qy       1585 CAGCCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCACTAACTGATCAAACTAACCGCAAA 1644
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         358 CAGCCAAGGAGAAGAAGAAGGTGGAAATGGCCGCCACTAACTGATCAAACTAACCGCAAA 299

Qy       1645 ATCACCAAATCCTAAGGGTTTTCTTGCAAAAACGTGTGCTGTACTGGCTAGCTAGCAAGT 1704
             ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db         298 ATCACCAAATCCTAAGGGTTTTCTTGCGAAAACGTGTGCTGTACTGGCTAGCTAGCAAGT 239

Qy       1705 AGTAGCAGCAACGTGGGAAGATTCGCTGATCCGGCGTTGCTGGAGAGCGACGGCCGGCGA 1764
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         238 AGTAGCAGCAACGTGGGAAGATTCGCTGATCCGGCGTTGCTGGAGAGCGACGGCCGGCGA 179

Qy       1765 CGACAAAGCTGAGCTCCAGCTCGAGACTTCTTAAAATCATCTTCAAGTACATGGATTTTA 1824
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         178 CGACAAAGCTGAGCTCCAGCTCGAGACTTCTTAAAATCATCTTCAAGTACATGGATTTTA 119

Qy       1825 TTTTGCTCTTTGCTTTGTCCGTAAAAGTTGTACTATGCGATGAAGAATACCAGTATGTAG 1884
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         118 TTTTGCTCTTTGCTTTGTCCGTAAAAGTTGTACTATGCGATGAAGAATACCAGTATGTAG 59

Qy       1885 CAAGGCTGAGGTTGTGTGTAGCTACTAGAAGTGTCAGTCACGTTGTTCTTGTAAGAAA 1942
             ||||| ||||  |||||||||||||||||||||||| ||||||||| |||||||
Db           58 CAAGGTTGAGAGTGTGTGTAGCTACTAGAAGTGTCAGTCCCGTTGTTCTTNTAAGAAA 1
```

RESULT 14
CG260290/c
LOCUS CG260290 929 bp DNA linear GSS 25-AUG-2003
DEFINITION OG3DF74TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0780N04, genomic survey sequence.
ACCESSION CG260290
VERSION CG260290.1 GI:34169668
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 929)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
TITLE Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Other_GSSs: OG3DF74TH
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
FEATURES Location/Qualifiers
source 1. .929
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMa0780N04"
/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"
ORIGIN

Query Match 32.2%; Score 648.6; DB 29; Length 929;
Best Local Similarity 81.6%; Pred. No. 7.7e-51;
Matches 750; Conservative 0; Mismatches 169; Indels 0; Gaps 0;

```
Qy        635 CATCATGCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCAT 694
              | |||||||||||||||||||||||||| |||  ||||||||||||||||||||| | ||
Db        929 CGTCATGCTCGGCATCGGCGCGGCGCCGGCCGGCCTGCTCGCGCTCATGGTGCTCTGGAT 870

Qy        695 GCCGGAGTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCT 754
              ||| |||||||| |||||||| |||||||||||.||||||| ||||||| ||  |||||
Db        869 GCCCGAGTCGCCTCGGTGGCTTGTCATGAAGGGCCGCCTCGCCGACGCCAGGGCCGTGCT 810

Qy        755 GGAGAAGACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGC 814
              |||||||||||| | | || |||||||||||||| ||||| ||||| ||||| ||||  ||
Db        809 GGAGAAGACCTCGGCCTCGCCGGAGGAGGCCGCCGAGCGGCTGGCTGACATTAAGGTTGC 750

Qy        815 CGCCGGCATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGGAGCGG 874
              ||| || || ||  |||  |||||||| ||||||||  |||| ||||||     | |||||
Db        749 CGCGGGGATTCCGAAGGGCCTCGACGGAGACGTGGTCGCCGTACCCAAGGAGCGAAGCGG 690

Qy        875 AAACGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCAT 934
                  ||||| || |||||||| |||||||||| | ||||| || ||||| || || |||||
Db        689 CGGCGAGATGCAGGTGTGGAGGGAGCTCATCTTATCCCCAACACCGGCGATCCGACGCAT 630

Qy        935 CCTGCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGT 994
              |||||||| |   ||||| | |||||||||||||| || |  ||||   | ||||||||
Db        629 ACTGCTGTCGGCCGTCGGCCTGCACTTCTTCCAGCAGGCCTCTGGCAGCGATTCCGTCGT 570

Qy        995 CTTCTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGG 1054
              | | ||||||||   |||||||||||||| | ||| | ||  |||||||   |  || ||||
Db        569 CCTGTACAGCCCGACCGTGTTCAAGAGCGCGGGAATCACCGACGACAACAAGCTCCTGGG 510
```

```
Qy       1055 CACCACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCAT 1114
              |  ||| ||  || | || ||| |||||| ||| |||||| || |||| || ||| |  |
Db        509 CGTCACCTGCGCGGTGGGCGTCGCCAAGACGCTCTTCATCCTGGTGGCCACGTTCCTGCT 450

Qy       1115 CGACGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCT 1174
                ||| ||| |||||| |||||  ||| |||||  |||||||| || || ||  |||| ||
Db        449 GGACCGCGCCGGGCGTCGGCCTCTGCTGCTGACCAGCACGGGTGGAATGATTGTCTCGCT 390

Qy       1175 CATCGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTG 1234
              | |||| |||||  | || |||||||||||| || || ||||| |||||||||||| || |
Db        389 CGTCGGTCTCGGGACGGGCCTCACCGTCGTGGGGCACCACCCGGACGCCAAGATCCCGTC 330

Qy       1235 GGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCT 1294
              |||| |||  ||  ||||||| |||||||| ||||||||||||||||||||||||||||||
Db        329 GGCCGTCGCGCTGTGCATCGCGTCCACCCTGGCCTACGTCGCCTTCTTCTCCATCGGCCT 270

Qy       1295 TGGCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGG 1354
              |||||| |||||  | |||||||| |||||||| |||||||| ||||||||||||||||
Db        269 CGGCCCCGTCACGGGTGTGTACAACTCGGAGATATTCCCGCTGCAGGTGCGCGCGCTGGG 210

Qy       1355 CTGCTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCT 1414
              | | || ||||||| ||   |||||||| ||||||| ||| ||||||||||||||||
Db        209 ATTCGCGGTCGGCGTGGCGTGCAACCGCGTTACCAGCGCCGTTATCTCCATGACCTTCCT 150

Qy       1415 GTCGCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGC 1474
              ||| |||||||||| ||||||||||||||||||||||||||||||| ||||||||||
Db        149 GTCCCTGTCCAAGGGTATCACCATCGGCGGCAGCTTCTTCCTCTACTCGGGCATCGCCGC 90

Qy       1475 GCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGA 1534
              | |||  ||||| ||||||| ||| | ||| ||||||||||||||| || ||||||||||
Db         89 GGTCGGGTGGGTTTTCTTCTTCACGTGCCTTCCGGAGACCCGCGGTCGTACGCTGGAGGA 30

Qy       1535 GATGAGCAAGCTGTTCGGC 1553
              |||| ||||||||||||||
Db         29 GATGGGCAAGCTGTTCGGC 11
```

RESULT 15
CG064683/c
LOCUS CG064683 972 bp DNA linear GSS 19-AUG-2003
DEFINITION PUJCW07TD ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa0639B13, genomic survey sequence.
ACCESSION CG064683
VERSION CG064683.1 GI:33936863
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 972)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and Bennetzen,J.
TITLE Maize Genomics Consortium
JOURNAL Unpublished (2003)
COMMENT Other_GSSs: PUJCW07TB
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
FEATURES Location/Qualifiers
source 1. .972
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTa0639B13"

/clone_lib="ZM_0.6_1.0_KB"
/note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high CoT selected genomic DNA library"

ORIGIN

Query Match 31.5%; Score 635; DB 29; Length 972;
Best Local Similarity 79.1%; Pred. No. 1.3e-49;
Matches 768; Conservative 0; Mismatches 200; Indels 3; Gaps 1;

```
Qy        641 GCTCGGCATCGGCGCGGCGCCGTCCGTGCTGCTCGCGCTCATGGTGCTCGGCATGCCGGA 700
              ||||| ||| ||||| |  ||||||| ||||||||||| ||||| || ||||||| ||
Db        972 GCTCGCCATTGGCGCAGTTCCGTCCGGCCTGCTCGCGCTCCTGGTGTTCTGCATGCCCGA 913

Qy        701 GTCGCCGCGGTGGCTGGTCATGAAGGGACGCCTCGCGGACGCCAAGGTGGTGCTGGAGAA 760
              |||||| |||||||||||||  |||||| |||||||||||||||| ||  ||||| |||||
Db        912 GTCGCCTCGGTGGCTGGTCTGGAAGGGCCGCCTCGCGGACGCCAGGGCTGTGCTAGAGAA 853

Qy        761 GACCTCCGACACGGCGGAGGAGGCCGCGGAGCGCCTGGCCGACATCAAGGCCGCCGCCGG 820
              |||||| | |||| | |||||||||||| ||||| ||||||||||||||||||||| || ||
Db        852 GACCTCTGCCACGCCAGAGGAGGCCGCCGAGCGGCTGGCCGACATCAAGGCCGCGGCGGG 793

Qy        821 CATCCCTGAGGAGCTCGACGGCGACGTGGTGACCGTCCCCAAGAGAGGG---AGCGGAAA 877
              || ||  |||  |||||||| |||||||| ||||| ||||| |  | |    ||||
Db        792 GATTCCGAAGGGCCTCGACGGGGACGTGGTCACCGTACCCAACAAGGAGCAAGGCGGCGG 733

Qy        878 CGAGAAGCGGGTGTGGAAGGAGCTCATCCTGTCCCCGACCCCGGCCATGCGGCGCATCCT 937
              |||  || |||||||||| |||||||||||||||||||||||  | || |||||||||
Db        732 TGAGTTGCAGGTGTGGAAGCAGCTCATCCTGTCCCCGACCCCGGCGGTCCGACGCATCCT 673

Qy        938 GCTGTCCGGGATCGGCATCCACTTCTTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTT 997
              ||| || |   | ||  |||||||||||||||| || |  ||||   ||||||||||
Db        672 GCTCTCGGCCGTGGGTCTCCACTTCTTCCAGCAGGCTTCTGGCAGCGACTCCGTCGTCCA 613

Qy        998 CTACAGCCCTCTCGTGTTCAAGAGCCCCGGATTAACGAACGACAAACACTTCTTGGGCAC 1057
              |||||| | | | |||||||||||| | ||  | ||  |||||||  |  || |||||
Db        612 GTACAGCGCCCGCCTGTTCAAGAGCGCGGGGATCACCGACGACAACAAGCTCCTGGGCGT 553

Qy       1058 CACTTGGCCGTTCGGTGTCACCAAGAGGCTTTTCATCTTGTTGGCGACTTTCTTCATCGA 1117
              ||| ||  || | || || ||||||| | | |||||| || |||| || ||| |  | ||
Db        552 CACCTGCGCGGTGGGCGTGACCAAGACGTTCTTCATCCTGGTGGCCACGTTCCTGCTGGA 493

Qy       1118 CGGCGTCGGGCGGCGGCCGCTGTTGCTGGGCAGCACGGGCGGGATAATCCTCTCCCTCAT 1177
              | |||  ||||| ||||| ||| |||||  |||||||||||||| ||| |||| |||||
Db        492 CCGCGCGGGGCGTCGGCCTCTGCTGCTGATCAGCACGGGCGGGATGATCGTCTCGCTCAT 433

Qy       1178 CGGCCTCGGCGCCGGGCTCACCGTCGTCGGCCAGCACCCCGACGCCAAGATACCTTGGGC 1237
              | |||||||  | |||||||||||||  || || ||||| ||| ||||| |  | |||||
Db        432 CTGCCTCGGGTCGGGGCTCACCGTCGCGGGGCATCACCCGGACACCAAGGTCGCGTGGGC 373

Qy       1238 CATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTCGCCTTCTTCTCCATCGGCCTTGG 1297
              | ||| |||  ||||||| ||||||||  ||||| |||||||||||||||||||||| ||
Db        372 CGTCGCCCTGTGCATCGCGTCCACCCTGTCCTACATCGCCTTCTTCTCCATCGGCCTCGG 313

Qy       1298 CCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTG 1357
              |||||||||  | |||||||  ||||||  || ||||||||  ||||||||||||||||||
Db        312 GCCCATCACGGGCGTGTACACCTCGGAAATATTCCCGCTGCAGGTGCGCGCGCTGGGCTT 253

Qy       1358 CTCGCTCGGCGTCGCCGCCAACCGCGTCACCAGCGGCGTCATCTCCATGACCTTCCTGTC 1417
              | || | || || ||   |||||||||||||||| ||||||||||||||||||||||||
Db        252 CGCGGTGGGTGTGGCGAGCAACCGCGTCACCAGCGCCGTCATCTCCATGACCTTCCTGTC 193

Qy       1418 GCTGTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGCT 1477
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        192 CCTCTCCAAGGCCATCACCATCGGCGGCAGCTTCTTCCTCTACTCCGGCATCGCCGCGGT 133

Qy       1478 CGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGGCCGGACGCTGGAGGAGAT 1537
              ||| ||||| ||||||| ||| | |||||||||||| ||||||||||||||||||||||
Db        132 CGCTTGGGTTTTCTTCTTCACGTGCCTCCCGGAGACACGCGGCCGGACGCTGGAGGAGAT 73

Qy       1538 GAGCAAGCTGTTCGGCGACACGGCCGCCGCCTCGGAATCAGACGAGCCAGCCAAGGAGAA 1597
              | |||||||||||||    | | | | | | ||   ||| |   || |   |
Db          72 GGGCAAGCTGTTCGGCATGCCAGACACGGGCATGGCTGAAGAAGCAGAAGACGCCGCAGC 13
```

```
Qy       1598 GAAGAAGGTGG 1608
              ||| ||  ||
Db         12 CAAGGAGAAGG 2
```

Search completed: July 2, 2004, 08:27:21
Job time : 7796 secs

4/14/99

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 30, 2004, 18:24:42 ; Search time 59 Seconds
(without alignments)
2456.727 Million cell updates/sec

Title: US-10-051-902A-20
Perfect score: 2559
Sequence: 1 MASDELAKAVEPRKKGNVKY.....AEEAEDAAAKEKVVELPSSK 513

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2559	100.0	513	5	AAU97210	Aau97210 Corn Beta
2	2559	100.0	513	6	ABU08335	Abu08335 Corn suga
3	1906.5	74.5	529	5	AAU97214	Aau97214 Wheat sug
4	1906.5	74.5	529	6	ABU08339	Abu08339 Wheat sug
5	1872.5	73.2	510	5	AAU97211	Aau97211 Rice Beta
6	1872.5	73.2	510	6	ABU08336	Abu08336 Rice suga
7	1764.5	69.0	539	5	AAU97213	Aau97213 Wheat sug
8	1764.5	69.0	539	6	ABU08338	Abu08338 Wheat sug
9	1542	60.3	523	5	AAU97212	Aau97212 Soybean B
10	1542	60.3	523	6	ABU08337	Abu08337 Soybean s
11	1420	55.5	513	5	ABB09681	Abb09681 Amino aci
12	1316	51.4	491	3	AAG32072	Aag32072 Arabidops
13	1316	51.4	508	3	AAG32071	Aag32071 Arabidops
14	1245	48.7	466	3	AAG32073	Aag32073 Arabidops
15	772.5	30.2	333	6	ADA48320	Ada48320 Rice prot
16	678	26.5	487	6	ABG73334	Abg73334 Consensus
17	664.5	26.0	488	6	ABP98504	Abp98504 PFAM cons
18	664.5	26.0	488	6	AAE35306	Aae35306 Human sug
19	664.5	26.0	488	7	ADD22918	Add22918 Human sug
20	646.5	25.3	629	4	AAB66935	Aab66935 Human GLU
21	646.5	25.3	629	4	AAB66940	Aab66940 GLUTX2 co
22	637.5	24.9	648	5	ABG61548	Abg61548 Human tra
23	624.5	24.4	618	4	AAB66936	Aab66936 Rat GLUTX
24	611	23.9	446	6	ABU43281	Abu43281 Protein e
25	602	23.5	478	3	AAG15416	Aag15416 Arabidops
26	602	23.5	493	3	AAG15415	Aag15415 Arabidops
27	602	23.5	546	3	AAG15414	Aag15414 Arabidops
28	599	23.4	450	6	ABU33853	Abu33853 Protein e
29	597	23.3	582	3	AAG29528	Aag29528 Arabidops
30	595	23.3	555	3	AAG29529	Aag29529 Arabidops
31	593.5	23.2	480	4	AAM00107	Aam00107 Sugar tra
32	586.5	22.9	472	5	ABP52164	Abp52164 E. coli a
33	586.5	22.9	472	7	ADD37502	Add37502 E. coli a
34	585	22.9	491	4	AAG89949	Aag89949 C glutami
35	583	22.8	491	3	AAB12594	Aab12594 Brevibact
36	576	22.5	167	5	AAU97209	Aau97209 Portion o
37	576	22.5	167	6	ABU08334	Abu08334 Corn suga
38	574.5	22.5	551	2	AAY50799	Aay50799 Spinach g
39	568	22.2	502	6	ABU36940	Abu36940 Protein e
40	565	22.1	584	5	ABB05603	Abb05603 Yeast ino
41	551.5	21.6	489	4	ABB63511	Abb63511 Drosophil
42	551.5	21.6	517	2	AAY49633	Aay49633 Wheat hex
43	548	21.4	517	5	ABP66239	Abp66239 Bifidobac
44	547	21.4	464	5	ABP52163	Abp52163 E. coli g
45	547	21.4	464	6	ABU14994	Abu14994 Protein e

ALIGNMENTS

Appl. priority doc.

RESULT 1
AAU97210
ID AAU97210 standard; protein; 513 AA.
XX
AC AAU97210;
XX
DT 27-AUG-2002 (first entry)
XX
DE Corn Beta vulgaris-like sugar transport protein.
XX
KW Corn; Beta vulgaris-like sugar transport protein; carbohydrate transport;
KW grain filling; annual field crop; plant.
XX
OS Zea mays.
XX
PN US6383776-B1.
XX
PD 07-MAY-2002.
XX
PF 14-APR-1999; 99US-00291922.
XX
PR 24-APR-1998; 98US-0083044P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX
DR WPI; 2002-453364/48.
DR N-PSDB; ABK51971.
XX
PT New nucleic acid encoding plant sugar-transport proteins, useful for
PT preparing transgenic plants with altered carbohydrate distribution.
XX
PS Example 4; Fig 2; 54pp; English.
XX
CC The present invention relates to the isolation of plant polynucleotide
CC sequences encoding an Arabidopsis thaliana-like sugar transport protein
CC or Beta vulgaris-like sugar transport protein. The polynucleotide
CC sequences are useful for altering the level of sugar transport proteins
CC in plants, i.e. for control of carbohydrate transport and distribution in
CC plant cells, e.g. during grain filling of annual field crops (e.g. corn,
CC rice, soybeans, and wheat), and, for studying carbohydrate flows and
CC sugar transport. The polynucleotide sequences can also be used to isolate
CC cDNA sequences and genes that encode homologues of the new proteins. The
CC present sequence represents a corn Beta vulgaris-like sugar transport
CC protein
XX
SQ Sequence 513 AA;

```
  Query Match             100.0%;  Score 2559;  DB 5;  Length 513;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-249;
  Matches  513;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MASDELAKAVEPRKKGNVKYASICAILASMASVILGYDIGVMSGAAMYIKKDLNITDVQL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MASDELAKAVEPRKKGNVKYASICAILASMASVILGYDIGVMSGAAMYIKKDLNITDVQL 60

Qy         61 EILIGILSLYSLFGSFAGARTSDRIGRRLTVVFAAVIFFVGSLLMGFAVNYGMLMAGRFV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 EILIGILSLYSLFGSFAGARTSDRIGRRLTVVFAAVIFFVGSLLMGFAVNYGMLMAGRFV 120

Qy        121 AGVGVGYGGMIAPVYTAEISPAASRGFLTTFPEVFINIGILLGYLSNFAFARLPLHLGWR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGVGVGYGGMIAPVYTAEISPAASRGFLTTFPEVFINIGILLGYLSNFAFARLPLHLGWR 180

Qy        181 VMLAIGAVPSGLLALLVFCMPESPRWLVLKGRLADARAVLEKTSATPEEAAERLADIKAA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VMLAIGAVPSGLLALLVFCMPESPRWLVLKGRLADARAVLEKTSATPEEAAERLADIKAA 240

Qy        241 AGIPKGLDGDVVTVPGKEQGGGELQVWKKLILSPTPAVRRILLSAVGLHFFQQASGSDSV 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGIPKGLDGDVVTVPGKEQGGGELQVWKKLILSPTPAVRRILLSAVGLHFFQQASGSDSV 300

Qy        301 VQYSARLFKSAGITDDNKLLGVTCAVGVTKTFFILVATFLLDRAGRRPLLLISTGGMIVS 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 VQYSARLFKSAGITDDNKLLGVTCAVGVTKTFFILVATFLLDRAGRRPLLLISTGGMIVS 360

Qy        361 LICLGSGLTVAGHHPDTKVAWAVALCIASTLSYIAFFSIGLGPITGVYTSEIFPLQVRAL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LICLGSGLTVAGHHPDTKVAWAVALCIASTLSYIAFFSIGLGPITGVYTSEIFPLQVRAL 420

Qy        421 GFAVGVASNRVTSAVISMTFLSLSKAITIGGSFFLYSGIAAVAWVFFFTCLPETRGRTLE 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GFAVGVASNRVTSAVISMTFLSLSKAITIGGSFFLYSGIAAVAWVFFFTCLPETRGRTLE 480

Qy        481 EMGKLFGMPDTGMAEEAEDAAAKEKVVELPSSK 513
              |||||||||||||||||||||||||||||||||
Db        481 EMGKLFGMPDTGMAEEAEDAAAKEKVVELPSSK 513


RESULT 2
ABU08335
ID   ABU08335 standard; protein; 513 AA.
XX
AC   ABU08335;
XX
DT   29-MAY-2003  (first entry)
XX
DE   Corn sugar transport protein #3.
XX
KW   Beta vulgaris-like sugar transport protein; corn; rice; wheat;
KW   plant sugar transport protein; carbohydrate transport; soybean;
KW   carbohydrate distribution; plant.
XX
OS   Zea mays.
XX
PN   US2002178468-A1.
XX
PD   28-NOV-2002.
XX
PF   17-JAN-2002; 2002US-00051902.
XX
PR   24-APR-1998;   98US-0083044P.
PR   14-APR-1999;   99US-00291922.
XX
PA   (ALLE/) ALLEN S M.
PA   (HITZ/) HITZ W D.
PA   (KINN/) KINNEY A J.
PA   (TING/) TINGEY S V.
XX
PI   Allen SM,  Hitz WD,  Kinney AJ,  Tingey SV;
XX
DR   WPI; 2003-340957/32.
DR   N-PSDB; ABX93207.
XX
PT   Novel plant sugar transport proteins and nucleic acid encoding the
PT   protein useful for producing transgenic plants having altered levels of
PT   sugar transport protein.
XX
PS   Claim 10; Fig 2; 56pp; English.
XX
CC   The present invention relates to the isolation of Arabidopsis thaliana-
CC   like or Beta vulgaris-like sugar transport proteins, and the
CC   polynucleotide sequences encoding them. The plant sugar transport
CC   proteins of the invention have been isolated from corn, rice, soybean,
CC   and wheat. The polypeptides of the invention may be used for altering the
CC   level of expression of a sugar transport protein in a host cell, by
CC   transforming a host cell with a chimeric construct encoding all, or a
CC   portion of the sugar transport protein, in sense or antisense
CC   orientation. Particularly, the polypeptides may provide a means to
CC   control carbohydrate transport and distribution in plants. ABU08334-
CC   ABU08339 represent Beta vulgaris-like sugar transport proteins
XX
SQ   Sequence 513 AA;

  Query Match             100.0%;  Score 2559;  DB 6;  Length 513;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-249;
  Matches  513;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MASDELAKAVEPRKKGNVKYASICAILASMASVILGYDIGVMSGAAMYIKKDLNITDVQL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MASDELAKAVEPRKKGNVKYASICAILASMASVILGYDIGVMSGAAMYIKKDLNITDVQL 60

Qy         61 EILIGILSLYSLFGSFAGARTSDRIGRRLTVVFAAVIFFVGSLLMGFAVNYGMLMAGRFV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 EILIGILSLYSLFGSFAGARTSDRIGRRLTVVFAAVIFFVGSLLMGFAVNYGMLMAGRFV 120

Qy        121 AGVGVGYGGMIAPVYTAEISPAASRGFLTTFPEVFINIGILLGYLSNFAFARLPLHLGWR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGVGVGYGGMIAPVYTAEISPAASRGFLTTFPEVFINIGILLGYLSNFAFARLPLHLGWR 180

Qy        181 VMLAIGAVPSGLLALLVFCMPESPRWLVLKGRLADARAVLEKTSATPEEAAERLADIKAA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VMLAIGAVPSGLLALLVFCMPESPRWLVLKGRLADARAVLEKTSATPEEAAERLADIKAA 240

Qy        241 AGIPKGLDGDVVTVPGKEQGGGELQVWKKLILSPTPAVRRILLSAVGLHFFQQASGSDSV 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGIPKGLDGDVVTVPGKEQGGGELQVWKKLILSPTPAVRRILLSAVGLHFFQQASGSDSV 300

Qy        301 VQYSARLFKSAGITDDNKLLGVTCAVGVTKTFFILVATFLLDRAGRRPLLLISTGGMIVS 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 VQYSARLFKSAGITDDNKLLGVTCAVGVTKTFFILVATFLLDRAGRRPLLLISTGGMIVS 360

Qy        361 LICLGSGLTVAGHHPDTKVAWAVALCIASTLSYIAFFSIGLGPITGVYTSEIFPLQVRAL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LICLGSGLTVAGHHPDTKVAWAVALCIASTLSYIAFFSIGLGPITGVYTSEIFPLQVRAL 420

Qy        421 GFAVGVASNRVTSAVISMTFLSLSKAITIGGSFFLYSGIAAVAWVFFFTCLPETRGRTLE 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GFAVGVASNRVTSAVISMTFLSLSKAITIGGSFFLYSGIAAVAWVFFFTCLPETRGRTLE 480

Qy        481 EMGKLFGMPDTGMAEEAEDAAAKEKVVELPSSK 513
              |||||||||||||||||||||||||||||||||
Db        481 EMGKLFGMPDTGMAEEAEDAAAKEKVVELPSSK 513


RESULT 3
AAU97214
ID   AAU97214 standard; protein; 529 AA.
XX
AC   AAU97214;
XX
DT   27-AUG-2002  (first entry)
XX
```

```
DE   Wheat sugar transport protein encoded by wlm1.pk0012.h1.
XX
KW   Wheat; Beta vulgaris-like sugar transport protein;
KW   carbohydrate transport; grain filling; annual field crop; plant.
XX
OS   Triticum aestivum.
XX
PN   US6383776-B1.
XX
PD   07-MAY-2002.
XX
PF   14-APR-1999;   99US-00291922.
XX
PR   24-APR-1998;   98US-0083044P.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Allen SM,  Hitz WD,  Kinney AJ,  Tingey SV;
XX
DR   WPI; 2002-453364/48.
DR   N-PSDB; ABK51975.
XX
PT   New nucleic acid encoding plant sugar-transport proteins, useful for
PT   preparing transgenic plants with altered carbohydrate distribution.
XX
PS   Example 4; Col 77-80; 54pp; English.
XX
CC   The present invention relates to the isolation of plant polynucleotide
CC   sequences encoding an Arabidopsis thaliana-like sugar transport protein
CC   or Beta vulgaris-like sugar transport protein. The polynucleotide
CC   sequences are useful for altering the level of sugar transport proteins
CC   in plants, i.e. for control of carbohydrate transport and distribution in
CC   plant cells, e.g. during grain filling of annual field crops (e.g. corn,
CC   rice, soybeans, and wheat), and, for studying carbohydrate flows and
CC   sugar transport. The polynucleotide sequences can also be used to isolate
CC   cDNA sequences and genes that encode homologues of the new proteins. The
CC   present sequence represents a wheat Beta vulgaris-like sugar transport
CC   protein
XX
SQ   Sequence 529 AA;

  Query Match             74.5%;  Score 1906.5;  DB 5;  Length 529;
  Best Local Similarity   73.6%;  Pred. No. 2.7e-183;
  Matches  373;  Conservative   58;  Mismatches   71;  Indels    5;  Gaps    3;

Qy          1 MASDELAK--AVEPRKKGNVKYASICAILASMASVILGYDIGVMSGAAMYIKKDLNITDV 58
              |||  | :  || || ||| |||  ||: ||||:::||||:||||||::|||:|| ||||
Db         19 MASAALPEPGAVHPRNKGNFKYAFTCALCASMATIVLGYDVGVMSGASLYIKRDLQITDV 78

Qy         59 QLEILIGILSLYSLFGSFAGARTSDRIGRRLTVVFAAVIFFVGSLLMGFAVNYGMLMAGR 118
              ||||::||||:|:| ||| |||||| :|||:|||||| ||  |||||||||||| ||| ||
Db         79 QLEIMMGILSVYALIGSFLGARTSDWVGRRVTVVFAAAIFNNGSLLMGFAVNYAMLMVGR 138

Qy        119 FVAGVGVGYGGMIAPVYTAEISPAASRGFLTTFPEVFINIGILLGYLSNFAFARLPLHLG 178
              || |:|||| |:||||| |:|||::|||||:| |||||:||||||:||:|||||||||
Db        139 FVTGIGVGYAIMVAPVYTPEVSPASARGFLTSFTEVFINVGILLGYVSNYAFARLPLHLS 198

Qy        179 WRVMLAIGAVPSGLLALLVFCMPESPRWLVLKGRLADARAVLEKTSATPEEAAERLADIK 238
              ||||| |||||| ||||:|| |||||||||:|||||||||| ||| |||| ||| ||
Db        199 WRVMLGIGAVPSALLALMVFGMPESPRWLVMKGRLADARAVLAKTSDTPEEAVERLDQIK 258

Qy        239 AAAGIPKGLDGDVVTVPGKEQGGGELQVWKKLILSPTPAVRRILLSAVGLHFFQQASGSD 298
              ||||||: |||||| :| | :|| | ||||:|| |||||:|||||:|:|:||||||:|||
Db        259 AAAGIPRELDGDVVVMP-KTKGGQEKQVWKELIFSPTPAMRRILLAALGIHFFQQATGSD 317

Qy        299 SVVQYSARLFKSAGITDDNKLLGVTCAVGVTKTFFILVATFLLDRAGRRPLLLISTGGMI 358
              ||| || |:|:||||| || ||| |||:|| || ||||||| ||| ||||||| || ||:
Db        318 SVVLYSPRVFQSAGITGDNHLLGATCAMGVMKTLFILVATFQLDRVGRRPLLLTSTAGML 377

Qy        359 VSLICLGSGLTVAGHHPDTKVAWAVALCIASTLSYIAFFSIGLGPITGVYTSEIFPLQVR 418
                || ||:|||| | ||| || ||: ||| | |:|::||||||||:| |||||:|||:||
Db        378 ACLIGLGTGLTVVGRHPDAKVPWAIGLCIVSILAYVSFFSIGLGPLTSVYTSEVFPLRVR 437
```

```
Qy        419 ALGFAVGVASNRVTSAVISMTFLSLSKAITIGGSFFLYSGIAAVAWVFFFTCLPETRGRT 478
              |||||:| : |||||| :||:|||||||||||||||||:||||: |:|||| :|||||
Db        438 ALGFALGTSCNRVTSAAVSMSFLSLSKAITIGGSFFLYAGIAAIGWIFFFTFIPETRGLP 497

Qy        479 LEEMGKLFGMPDTGMAEEAEDAAAKEK 505
              |||:|||||| ||  | ||:| | |:|
Db        498 LEEIGKLFGMTDT--AVEAQDTATKDK 522


RESULT 4
ABU08339
ID   ABU08339 standard; protein; 529 AA.
XX
AC   ABU08339;
XX
DT   29-MAY-2003  (first entry)
XX
DE   Wheat sugar transport protein #5.
XX
KW   Beta vulgaris-like sugar transport protein; corn; rice; wheat;
KW   plant sugar transport protein; carbohydrate transport; soybean;
KW   carbohydrate distribution; plant.
XX
OS   Triticum aestivum.
XX
PN   US2002178468-A1.
XX
PD   28-NOV-2002.
XX
PF   17-JAN-2002; 2002US-00051902.
XX
PR   24-APR-1998;   98US-0083044P.
PR   14-APR-1999;   99US-00291922.
XX
PA   (ALLE/) ALLEN S M.
PA   (HITZ/) HITZ W D.
PA   (KINN/) KINNEY A J.
PA   (TING/) TINGEY S V.
XX
PI   Allen SM,  Hitz WD,  Kinney AJ,  Tingey SV;
XX
DR   WPI; 2003-340957/32.
DR   N-PSDB; ABX93211.
XX
PT   Novel plant sugar transport proteins and nucleic acid encoding the
PT   protein useful for producing transgenic plants having altered levels of
PT   sugar transport protein.
XX
PS   Claim 10; Fig 2; 56pp; English.
XX
CC   The present invention relates to the isolation of Arabidopsis thaliana-
CC   like or Beta vulgaris-like sugar transport proteins, and the
CC   polynucleotide sequences encoding them. The plant sugar transport
CC   proteins of the invention have been isolated from corn, rice, soybean,
CC   and wheat. The polypeptides of the invention may be used for altering the
CC   level of expression of a sugar transport protein in a host cell, by
CC   transforming a host cell with a chimeric construct encoding all, or a
CC   portion of the sugar transport protein, in sense or antisense
CC   orientation. Particularly, the polypeptides may provide a means to
CC   control carbohydrate transport and distribution in plants. ABU08334-
CC   ABU08339 represent Beta vulgaris-like sugar transport proteins
XX
SQ   Sequence 529 AA;

  Query Match             74.5%;  Score 1906.5;  DB 6;  Length 529;
  Best Local Similarity   73.6%;  Pred. No. 2.7e-183;
  Matches  373;  Conservative   58;  Mismatches   71;  Indels    5;  Gaps    3;

Qy          1 MASDELAK--AVEPRKKGNVKYASICAILASMASVILGYDIGVMSGAAMYIKKDLNITDV 58
              |||  | :  || || ||| |||  ||: ||||:::||||:||||||::|||:|| ||||
Db         19 MASAALPEPGAVHPRNKGNFKYAFTCALCASMATIVLGYDVGVMSGASLYIKRDLQITDV 78

Qy         59 QLEILIGILSLYSLFGSFAGARTSDRIGRRLTVVFAAVIFFVGSLLMGFAVNYGMLMAGR 118
```

```
Db         79 QLEIMMGILSVYALIGSFLGARTSDWVGRRVTVVFAAAIFNNGSLLMGFAVNYAMLMVGR 138

Qy        119 FVAGVGVGYGGMIAPVYTAEISPAASRGFLTTFPEVFINIGILLGYLSNFAFARLPLHLG 178

Db        139 FVTGIGVGYAIMVAPVYTPEVSPASARGFLTSFTEVFINVGILLGYVSNYAFARLPLHLS 198

Qy        179 WRVMLAIGAVPSGLLALLVFCMPESPRWLVLKGRLADARAVLEKTSATPEEAAERLADIK 238

Db        199 WRVMLGIGAVPSALLALMVFGMPESPRWLVMKGRLADARAVLAKTSDTPEEAVERLDQIK 258

Qy        239 AAAGIPKGLDGDVVTVPGKEQGGGELQVWKKLILSPTPAVRRILLSAVGLHFFQQASGSD 298

Db        259 AAAGIPRELDGDVVVMP-KTKGGQEKQVWKELIFSPTPAMRRILLAALGIHFFQQATGSD 317

Qy        299 SVVQYSARLFKSAGITDDNKLLGVTCAVGVTKTFFILVATFLLDRAGRRPLLLISTGGMI 358

Db        318 SVVLYSPRVFQSAGITGDNHLLGATCAMGVMKTLFILVATFQLDRVGRRPLLLTSTAGML 377

Qy        359 VSLICLGSGLTVAGHHPDTKVAWAVALCIASTLSYIAFFSIGLGPITGVYTSEIFPLQVR 418

Db        378 ACLIGLGTGLTVVGRHPDAKVPWAIGLCIVSILAYVSFFSIGLGPLTSVYTSEVFPLRVR 437

Qy        419 ALGFAVGVASNRVTSAVISMTFLSLSKAITIGGSFFLYSGIAAVAWVFFFTCLPETRGRT 478

Db        438 ALGFALGTSCNRVTSAAVSMSFLSLSKAITIGGSFFLYAGIAAIGWIFFFTFIPETRGLP 497

Qy        479 LEEMGKLFGMPDTGMAEEAEDAAAKEK 505

Db        498 LEEIGKLFGMTDT--AVEAQDTATKDK 522


RESULT 5
AAU97211
ID   AAU97211 standard; protein; 510 AA.
XX
AC   AAU97211;
XX
DT   27-AUG-2002  (first entry)
XX
DE   Rice Beta vulgaris-like sugar transport protein.
XX
KW   Rice; Beta vulgaris-like sugar transport protein; carbohydrate transport;
KW   grain filling; annual field crop; plant.
XX
OS   Oryza sativa.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 102
FT                   /label= Unknown
XX
PN   US6383776-B1.
XX
PD   07-MAY-2002.
XX
PF   14-APR-1999;   99US-00291922.
XX
PR   24-APR-1998;   98US-0083044P.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Allen SM,  Hitz WD,  Kinney AJ,  Tingey SV;
XX
DR   WPI; 2002-453364/48.
DR   N-PSDB; ABK51972.
XX
PT   New nucleic acid encoding plant sugar-transport proteins, useful for
PT   preparing transgenic plants with altered carbohydrate distribution.
XX
PS   Example 4; Fig 2; 54pp; English.
XX
CC   The present invention relates to the isolation of plant polynucleotide
CC   sequences encoding an Arabidopsis thaliana-like sugar transport protein
CC   or Beta vulgaris-like sugar transport protein. The polynucleotide
CC   sequences are useful for altering the level of sugar transport proteins
CC   in plants, i.e. for control of carbohydrate transport and distribution in
CC   plant cells, e.g. during grain filling of annual field crops (e.g. corn,
CC   rice, soybeans, and wheat), and, for studying carbohydrate flows and
CC   sugar transport. The polynucleotide sequences can also be used to isolate
CC   cDNA sequences and genes that encode homologues of the new proteins. The
CC   present sequence represents a rice Beta vulgaris-like sugar transport
CC   protein
XX
SQ   Sequence 510 AA;

  Query Match             73.2%;  Score 1872.5;  DB 5;  Length 510;
  Best Local Similarity   72.7%;  Pred. No. 7.1e-180;
  Matches  372;  Conservative   60;  Mismatches   77;  Indels    3;  Gaps    2;

Qy          1 MASDELAKAVEPRKKGNVKYASICAILASMASVILGYDIGVMSGAAMYIKKDLNITDVQL 60

Db          1 MASAALPEAVAPKKKGNVRFAFACAILASMTSILLGYDIGVMSGASLYIKKDFNISDGKV 60

Qy         61 EILIGILSLYSLFGSFAGARTSDRIGRRLTVVFAAVIFFVGSLLMGFAVNYGMLMAGRFV 120

Db         61 EVLMGILNLYSLIGSFAAGRTSDWIGRRYTIVFAAVIFFAGXFLMGFAVNYAMLMFGRFV 120

Qy        121 AGVGVGYGGMIAPVYTAEISPAASRGFLTTFPEVFINIGILLGYLSNFAFARLPLHLGWR 180

Db        121 AGIGVGYALMIAPVYTAEVSPASARGFLTSFPEVFINFGILLGYVSNYAFSRLPLNLGWR 180

Qy        181 VMLAIGAVPSGLLALLVFCMPESPRWLVLKGRLADARAVLEKTSATPEEAAERLADIKAA 240

Db        181 IMLGIGAAPSVLLALMVLGMPESPRWLVMKGRLADAKVVLEKTSDTAEEAAERLADIKAA 240

Qy        241 AGIPKGLDGDVVTVPGKEQGGGELQVWKKLILSPTPAVRRILLSAVGLHFFQQASGSDSV 300

Db        241 AGIPEELDGDVVTVP-KRGSGNEKRVWKELILSPTPAMRRILLSGIGIHFFQHALGIHSV 299

Qy        301 VQYSARLFKSAGITDDNKLLGVTCAVGVTKTFFILVATFLLDRAGRRPLLLISTGGMIVS 360

Db        300 VFYSPLVFKSPGLTNDKHFLGTTWPFGVTKRLFILLATFFIDGVGRRPLLLGSTGGIILS 359

Qy        361 LICLGSGLTVAGHHPDTKVAWAVALCIASTLSYIAFFSIGLGPITGVYTSEIFPLQVRAL 420

Db        360 LIGLGAGLTVVGQHPDAKIPWAIGLSIASTLAYVAFFSIGLGPITWVYSSEIFPLQVRAL 419

Qy        421 GFAVGVASNRVTSAVISMTFLSLSKAITIGGSFFLYSGIAAVAWVFPFTCLPETRGRTLE 480

Db        420 GCSLGVAANRVTSGVISMTFLSLSKAITIGGSFFLYSGIAALAWVFFYTYLPETRGRTLE 479

Qy        481 EMGKLFGMPDTGMAEEAEDAAAKEKVVELPSS 512

Db        480 EMSKLFG--DTAAASESDEPAKEKKKVEMAAT 509


RESULT 6
ABU08336
ID   ABU08336 standard; protein; 510 AA.
XX
AC   ABU08336;
XX
DT   29-MAY-2003  (first entry)
XX
DE   Rice sugar transport protein #3.
XX
KW   Beta vulgaris-like sugar transport protein; corn; rice; wheat;
KW   plant sugar transport protein; carbohydrate transport; soybean;
KW   carbohydrate distribution; plant.
XX
OS   Oryza sativa.
XX
PN   US2002178468-A1.
XX
PD   28-NOV-2002.
XX
```